

This is a repository copy of *Semantic indeterminacy in object relative clauses*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/3724/>

Article:

Gennari, Silvia P. orcid.org/0000-0002-2242-4002 and MacDonald, Maryellen C. (2008) Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*. pp. 161-187. ISSN 0749-596X

<https://doi.org/10.1016/j.jml.2007.07.004>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

promoting access to White Rose research papers



Universities of Leeds, Sheffield and York
<http://eprints.whiterose.ac.uk/>

This is an author produced version of a paper published in **Journal of Memory and Language**.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/3724/>

Published paper

Gennari, S. and MacDonald, M.C. (2008) *Semantic indeterminacy in object relative clauses*, Journal of Memory and Language, Volume 58 (2), 161 -187.

Semantic indeterminacy in object relative clauses

Silvia P. Gennari

University of York

Maryellen C. MacDonald

University of Wisconsin-Madison

Corresponding author

Silvia P. Gennari

University of York

Department of Psychology

Heslington, York, YO10 5DD

United Kingdom

Phone: +44 (0) 1904 432877

Fax: +44 (0)1904 433181

Abstract

This article examined whether semantic indeterminacy plays a role in comprehension of complex structures such as object relative clauses. Study 1 used a gated sentence completion task to assess which alternative interpretations are dominant as the relative clause unfolds; Study 2 compared reading times in object relative clauses containing different animacy configurations to unambiguous passive controls; and Study 3 related completion data and reading data. The results showed that comprehension difficulty was modulated by animacy configuration and voice (active vs. passive). These differences were well correlated with the availability of alternative interpretations as the relative clause unfolds, as revealed by the completion data. In contrast to approaches arguing that comprehension difficulty stems from syntactic complexity, these results suggest that semantic indeterminacy is a major source of comprehension difficulty in object relative clauses. Results are consistent with constraint-based approaches to ambiguity resolution and bring new insights into previously identified sources of difficulty.

Keywords: reading comprehension, object relative clauses, thematic roles.

Introduction

Research in sentence comprehension has pursued two distinct paths, one addressing the question of how comprehenders resolve temporary syntactic ambiguities and the other investigating syntactically complex but apparently unambiguous structures. Examples of temporary ambiguity are sentences containing reduced relatives such as *The man examined by the doctor ...*, whereas sentences containing center embedded or object relative clauses, such as *The man that the doctor examined* have been the central example of the latter category. This division has had a profound effect on the kinds of accounts that have been proposed for comprehension processes in these two cases.

The ambiguity resolution literature has centered on the debate between two-stage vs. constraint-based approaches to comprehension processes in ambiguous constructions. Two-stage models (e.g. Frazier & Fodor, 1978; Frazier & Clifton, 1989; Clifton & Frazier, 1989; Frazier & Rayner, 1982; Pickering & Traxler, 1998) proposed that comprehenders initially entertained only one analysis for an ambiguous structure by following simple parsing principles. This initial parsing triggered a process of reanalysis if later incoming information signaled that the initially postulated structure was incorrect. Comprehension difficulty in this account thus stemmed from the complexity of the reanalysis process. Constraint-based accounts of ambiguity resolution, in contrast, proposed that alternative interpretations are partially activated as a function of their frequency, plausibility, and other constraints (MacDonald, Pearlmutter & Seidenberg, 1994; MacRae, Spivey-Knowlton & Tanenhaus, 1998; Tanenhaus & Trueswell, 1995; Trueswell, Tanenhaus & Garnsey, 1994). On this view, the difficulty observed in interpreting ambiguous structures can be traced to competition between alternative interpretations.

The syntactic complexity literature, which we review in detail below, has instead focused on memory demands, thematic role assignments and other hypothesized sources of comprehension

difficulty in processing complex structures (e.g., Gibson, 1998; MacWhinney & Pleh, 1988; Waters & Caplan, 1996a,b). One prevailing view within this literature is that key mechanisms of processing operating in ambiguity resolution have little influence in comprehension of complex unambiguous structures such as object relative clauses. Gibson (1998), for example, argued that complexity effects found in comparing structures of varying complexity are not caused by differences in ambiguity, and consequently, neither two-stage nor constraint based theories of syntactic ambiguity resolution make any predictions in these cases (see p. 2). More recently, Gordon, Hendrick, and Johnson (2004) and Grodner and Gibson (2005) have endorsed a similar view and have argued that frequency information, a major constraint in ambiguity resolution, could not account for processing difficulty in object relative clauses.

One notable exception to this general division between the ambiguity and the complexity literature is an account by Traxler, Morris and Seely (2002) who proposed that complexity effects in object relative clauses are in fact due to a two-stage ambiguity resolution process of the sort proposed in the ambiguity literature. This move thus challenged the basic assumption that object relative clauses had no significant ambiguity and instated an ambiguity resolution theory within the complexity literature. Traxler et al. argued that a local temporary indeterminacy at the relative pronoun *that* is the major source of comprehension difficulty in object relative clauses, as for example, the noun phrase *the man* in *the man that* can either be the subject or object of the upcoming relative clause. This local indeterminacy had been previously noted (Gibson, 1998) but has not been considered a major source of comprehension difficulty because the very next word eliminates the indeterminacy, e.g. in *the man that the...* the head noun cannot be interpreted as the subject of the upcoming relative. Thus Traxler et al.'s account has not had the effect of linking the sentence complexity and ambiguity resolution literature, both because many

consider the ambiguity at *that* a trivial one and also because Traxler et al. did not address constraint-based approaches to ambiguity resolution.

In this paper, we attempt to fill the gap between complexity and ambiguity resolution accounts by investigating whether probabilistic approaches are truly irrelevant to the vast research on object relative clauses. We ask (a) whether object relative clauses display parsing and semantic indeterminacy as they unfold over time, analogous to more traditional syntactic ambiguities, and (b) whether the activation of various competing interpretations (indeterminacy) can account for comprehension difficulty. To foreshadow our results, we do find evidence for significant semantic indeterminacy in object relatives, thus suggesting that constraint based accounts of ambiguity resolution can offer insight into the difficulties associated with interpretation of object relative clauses, a domain previously thought to be outside the range of this approach.

Processing Difficulty in Object Relative Clauses

In investigating the comprehension difficulty presented by object relative clauses, researchers have frequently contrasted them with subject relative clauses, as in examples (1a-b) used by King and Just (1991).

(1) a. **Subject relative** The reporter that attacked the senator admitted the error.

b. **Object relative** The reporter that the senator attacked admitted the error.

The distinction between subject and object relatives refers to the fact that the modified noun phrase *the reporter* serves different roles of the verb of the subordinate clause *attacked* – subject in (1a) and object in (1b). A large number of studies have documented that object relative clauses are more difficult than subject relatives by measures such as reading times, often accompanied by error rates in comprehension questions after reading (Gordon, Hendrick, & Johnson, 2001; 2004; Just & Carpenter, 1992; King & Just, 1991; MacWhinney & Pleh, 1988, Mak, Vonk &

Schriefers, 2002, 2006; Traxler, et al, 2002, Waters & Caplan, 1996; 1996; Warren & Gibson, 2002), and performance errors in tasks such as lexical decision and word recall taking place during or after reading (Ford, 1983; Baird & Koslick, 1974; Waters et al., 1987, Waters & Caplan, 1999).

The choice of subject relatives as a baseline condition from which to observe comprehension difficulty in object relatives has partially determined the kinds of accounts that have been proposed. The words in structures such as (1a) and (1b) may be identical, but the two sentence types differ in their word order and in final sentence meaning. This choice contrasts with typical baseline conditions in ambiguity resolution studies in which unambiguous structures are chosen to have similar meanings to those in ambiguous conditions, though the sentences may have minor differences in number of words or word choices. Because subject and object relatives differ in both word order and final sentence meaning, it is possible to attribute differences in comprehension difficulty to either of these factors, and both sorts of accounts have been proposed in the literature (though they have not been identified as such). For example, accounts offered by Sheldon (1974), MacWhinney (1977; 1982) and MacWhinney and Pleh (1988) can be seen as supporting the view that the final meaning of the object relative is more complex than that of its subject relative counterpart. In object relatives, the modified head noun (e.g. *reporter* in (1b)) is simultaneously the patient of the relative clause verb (*attacked*) and the agent of the main verb (*admitted*). By contrast, subject relatives such as the example in (1a) have what is taken to be a simpler meaning, in which the head noun is the agent of both the embedded verb (*attacked*) and the main verb (*admitted*). On this view, comprehension difficulty for object relatives can be traced to the extra burden associated with assigning conflicting thematic roles to the head noun (what MacWhinney & Pleh, 1988, called perspective shifting).

Other accounts trace difficulty in relative clause processing to parsing differences that stem from encountering the different word orders in the two sentence types. A dominant theme in this work is that object relatives are difficult because comprehenders encounter two un-integrated nouns (e.g. *The reporter that the senator*) before a verb, thus creating a working memory load in which two nouns must be held in memory while waiting for verb information that relates them (Gibson, 1998; King & Just, 1991; Wanner & Maratsos, 1978, Caplan & Waters, 1999, Waters et al. 1987, Waters & Caplan 1996). By contrast, in subject relatives such as (1a), the alternating pattern of nouns and verbs permits rapid integration of these sentence elements, reducing the memory burden from maintaining un-integrated constituents.

Recent developments in these memory-based accounts have incorporated semantic factors into their explanations in order to accommodate the fact that comprehension difficulty of object relatives varies with the referential properties of the nouns they contain. Indexical expressions such as *I* or *you* and proper names in the embedded subject position (e.g. *The reporter that I attacked admitted the error*) reduce or eliminate the difficulty of object relatives compared to subject relatives (Gordon et al., 2001) and are less difficult than other object relatives with descriptive noun phrases such as *the lawyer* (Warren & Gibson, 2002; Gordon et al., 2001). These findings were taken to indicate that (a) the greater semantic availability of the indexical referents facilitates their access and integration into the sentential representation, by easing memory load (Warren & Gibson, 2002), or (b) semantically similar descriptive noun phrases (as in *the reporter that the senator ...*) compete and interfere with each other in memory (Gordon et al., 2001, 2004), thus increasing memory load and further integration (for a different view, see Real and Christiansen, 2007).

An approach of a different nature has emerged from recent findings deemed incompatible with the memory-based accounts. Mak et al. (2002, 2006), Traxler et al. (2002) and Traxler,

Williams, Blozis & Morris (2005) have observed that comprehension difficulty in object relatives varies with the animacy of the nouns involved. In these studies, the animacy configuration exemplified in (2d) below was almost as easily comprehended as subject relatives with either animate or inanimate heads in examples (2a) and (2c), whereas object relatives with animate heads were more difficult than any other structure in (2) including the object relatives in (2d).

- (2) a. The director that watched the movie received a prize. (Subj. Rel. - Animate head)
 b. The director that the movie pleased received a prize. (Obj. Rel. – Animate head)
 c. The movie that pleased the director received a prize. (Subj. Rel. – Inanimate head)
 d. The movie that the director watched received a prize. (Obj. Rel. – Inanimate head)

These results are a challenge for memory based approaches because interference between the nouns and their semantic accessibility remained constant across conditions. In arguing against the memory-interference approach, Mak et al. (2006) proposed that relative clause processing is determined by the interplay of factors such as topichood and animacy. Traxler et al. (2002) in contrast argued that comprehension difficulty could be traced to the subject relative vs. object relative structural ambiguity arising at *that*. On the two-stage ambiguity resolution approach, comprehenders initially interpret sentences such as (2a-d) using a structural parsing heuristic called the active filler strategy (Clifton & Frazier, 1989; Frazier & Clifton 1989; Pickering & Traxler 1998; Traxler & Pickering 1996) according to which the parser assumes a syntactic position for the antecedent of a pronoun as soon as possible within the unfolding structure. Use of this strategy forces interpretation of the head noun (e.g. *the director* in 2b) as the subject of the relative clause verb (e.g., *pleased*), leading to misanalyses of object relative clauses (2b, 2d) as subject relatives (2a, 2c). Later information within the relative clause would indicate that this initial analysis is incorrect, triggering a process of reanalysis in a second stage.

In this view, animacy considerations come into play at reanalysis: animate heads are more difficult to reanalyze because they are good syntactic subjects and agents, causing more delay in abandoning the initial subject analysis.

The constraint based approach

The animacy effects found in processing object relatives are interesting because they point to a source of processing difficulty already studied in the ambiguity resolution literature within a constraint-based framework. Trueswell et al. (1994), for example, showed that noun animacy and its semantic fit with the verb's thematic roles had immediate effects on processing difficulty (see also Mc Rae, et al. 1998; Tabor and Tanenhaus, 1999; Tanenhaus, Carlson & Trueswell, 1989). In main verb/reduced relative ambiguities such as *the man examined by the doctor* (where *the man* is initially taken as the agent of *examined*), the presence of an inanimate head (e.g., *the evidence examined*) eliminates the processing cost associated with the ambiguity. Inanimate nouns are typically theme-arguments of verbs and they are quickly processed as such. Animate nouns, in contrast, are typically correlated with agent-arguments and tend to occur in main clauses. Consequently, comprehenders require the disambiguating *by*-phrase in *the man examined by the doctor* to abandon the initial analysis of the noun as agent of a main clause. Thus, although structural ambiguity (that is, indeterminacy in assigning syntactic structure to an input string) accounts for some processing difficulty in ambiguity resolution, the indeterminacy of thematic role assignments (uncertainty as to who did what to whom in the sentence) is also a significant source of computational difficulty.

Both structural and semantic indeterminacies need not operate only in structurally ambiguous structures. In constraint-based approaches, syntactic and semantic structures are hypothesized to be continuously activated in parallel at each point in time (cf. Boland, 1997; Boland, Tanenhaus, Garnsey, & Carlson, 1995; MacDonald, et al, 1994; MacDonald, 1994;

MacRae et al, 1998; Spivey-Knowlton & Sedivy, 1995; Tanenhaus & Trueswell, 1995; Trueswell, et al, 1994). Independent of structural ambiguity, semantic and syntactic indeterminacies may emerge from the activation of several alternative structures. Comprehension difficulty in this approach emerges from competition between alternative structures and interpretations that are partially activated during comprehension. Which alternatives are entertained and their activation levels are ultimately determined by their frequency in speakers' linguistic experience.

These observations suggest that comprehension difficulty in object relatives may emerge from the activation of several competing structures ultimately derived from distributional patterns of language use such as noun animacy configurations and noun-verb co-occurrences (thematic fit). Object relative clauses with different animacy configurations may involve different competition processes between structural and semantic analyses, thus engendering diverging degrees of indeterminacy and difficulty. For instance, the nouns in a sequence such as *the director that the movie...* in (2b) can alternatively be analyzed as the agent and theme of a passive structure (e.g., *the director that the movie was written by*), as the goal and theme of a preposition verb phrase (e.g., *the director that the movie was for/was given to*) or as the themes of a transitive relational event (e.g., *the director that the movie depicted/was about*), none of which are consistent with the thematic roles of *please*, the upcoming verb in (2b). In contrast, the nouns in a structure such as *the movie that the director...* in (2d) are most likely interpreted as the theme and agent of an event (e.g., *wrote*). Infrequent and unexpected structures in a given configuration would then be difficult to activate because of competition with more available frequent structures. Thus, the relative ease or difficulty in processing relative clauses may depend on the extent of the competition and structural indeterminacy that each structure affords.

This approach is similar in spirit to several information theoretic perspectives on sentence processing, including *surprisal*, entropy reduction, and expectation-based frameworks (Hale, 2001; 2006; Levy 2007; Jurafsky, 2003). These approaches are not identical (see Levy, 2007, for discussion), but for the present purposes can be described informally as predicting comprehension difficulty as a function of uncertainty concerning upcoming material in the sentence, such that highly expected material is read more quickly than material that is less easy to predict. This approach has clear similarities with constraint-based processing (Levy, 2007), though particular formulations of the approach yield different predictions concerning the precise locus of processing difficulty. In particular, Levy (2007) suggested that relative clauses pose particular challenges for expectation-based accounts, presumably including both the information-theoretic and constraint-based variants. However, this view may reflect an incomplete assessment of the nature and loci of indeterminacy in these structures. We will consider these claims further in the general discussion.

To test the degree of indeterminacy in object relative clauses, we conducted several studies. First, we conducted a series of gated completions studies (McRae et al., 1998) to investigate the kind and the frequency of the structures that speakers entertain at several points within object relative clauses of varying noun animacy. The relative clause prompts used in the completions were based on the materials of Experiment 3 in Traxler et al. (2002), which contrasted the animacy configurations in (2b, 2d). We assume that the participants' preferences in this task reflect broad distributional preferences of language use: participants tend to provide completions that are consistent with frequent uses of the fragment prompts and the constructions that they themselves would be likely to use in their productions. Next, we conducted a self-paced reading experiment with the full version of relative clauses in Traxler et al. (2002), though we used a different baseline condition than the subject relatives used by Traxler et al. and typical of

most studies on relative clause processing. Finally, we regress the proportion of completions onto the reading times of our reading study to establish whether the interpretations entertained at each point in the gated completion studies predict reading times at various positions within the relative clause.

Study 1: Gated sentence completions

In this study, we investigated the interpretations that comprehenders entertain as the relative clause unfolds. To this end, we conducted a series of sentence completion tasks at different relative clause positions. We used materials from Traxler et al.'s (2002) object relative sentences to create the prompts from which completions were elicited. These materials had either an animate or an inanimate head noun as in (2b,d) (see Table 1 and Appendix A). For example, for animate and inanimate head relative clauses such as *the director that the movie pleased* and *the movie that the director watched*, we elicited completions at different positions such as *the director that...*, *the director that the...*, *the director that the movie...*. We also compared this last relative clause prompt to passive prompts containing the same number of words and headed by the same noun (*the director that was pleased...*). The passive relatives, which are a form of subject relative clause (as the head noun is the subject of the relative clause verb) are good semantic matches for object relatives because the noun animacy and the nouns' thematic roles are the same in both cases. Thus they provide a semantically-similar control condition to the structure of interest, object relatives, much the same way that unambiguous control conditions are similar in meaning to ambiguous structures in many ambiguity resolution studies. We hypothesized that noun animacy would play a critical role in determining the sort of interpretation that readers in the completion task entertain at various points in the relative clause because animacy is strongly correlated in speakers' experience with certain event roles (Trueswell et al., 1994; McRae et al., 1998). In addition, because animate nouns can play

different and more roles in an event than inanimate nouns, we expected that active animate-head object relatives would show more indeterminacy and more conflicting analyses than inanimate-head object relatives. Likewise, we expected active object relatives to differ from their passive counterparts: as active object relatives unfold, there should present more semantic indeterminacy because the verb, which defines the roles that participants play in the event, comes later in these structures compared to passive relative clauses. Thus, before the relative clause verb is encountered in active relatives there would be more competing structural analyses and interpretations assigned to the two relative clause nouns. In contrast, the early relative clause verb in passive relative clauses reduces the indeterminacy associated with the event role of the preceding noun and allows predicting or activating the upcoming material with more certainty (Altmann and Kamide, 1999).

To evaluate semantic indeterminacy, we coded the sentence completions according to their semantic characteristics such as the choice of noun animacy in the completions and the thematic roles assigned to the nouns by the verb chosen. This allowed us to compute the number and likelihood of the different interpretations entertained at various relative clause positions. One problem with this approach is that there is little agreement among linguists concerning which thematic roles exist and how to independently justify the use of a particular thematic role in a particular case (Dowty, 1991). Some linguistics introductory textbooks distinguish between patient and theme roles (the theme being the moving object in an event) (Haegeman, 1991), whereas others consider both the moving entity and the entity undergoing the action in an event to bear a theme role (O'Grady, Archibald, Aronoff and Rees-Miller, 2001). However, several semanticists have argued that thematic roles should not be conceived as conceptual or theoretical primitives. Rather they simply represent convenient labels to encompass clusters of properties that are shared across verbs or events of certain types (Dowty, 1991, Chierchia and McConnell-

Ginet, 1990, Grimshaw, 1990; Jackendoff, 1990; Ferretti, McRae and Hatherell, 2001). In this view, the lexical meaning of individual verbs (and more generally, the conceptual representation of the event that the verb refers to) carries with them clusters of entailments or properties concerning the arguments of the verbs (the event participants). For example, *kick* entails a participant doing the kicking (which in turn must have legs, be able to move, be sentient, etc.) and another participant to which the kicking is directed. In this view, when verb meanings require animate or inanimate participants, it is only appropriate that the thematic role labels assigned to them would capture these properties, as the animacy of an event participant is obviously correlated with the type of role it can play in the event. Indeed, many definitions of thematic roles include properties such as sentient, intentional or volitional participant (i.e., animate participant), as in the case of agents and experiencers (Dowty, 1989).

Method

Participants. A total of 64 students at the University of Wisconsin-Madison participated in this study for course credit. There were 16 participants for each of the 4 questionnaires. Each questionnaire in turn had 2 lists so that 8 participants filled each list.

Materials. We constructed sentence fragments from the materials in Appendix A, with animacy and relative clause voice as main grouping factors. Table 2 shows the relative clause positions at which sentence completions were elicited for both actives and passives. For active completions at position 1, the sentence fragment to be completed contained the relative clause head plus the relative clause pronoun (e.g., *the director that ...*), and the head noun was either animate or inanimate as in Table 1. At position 2, the fragment was extended to include the next word, the determiner *the* (e.g., *the director that the ...*). At position 3, the fragment included up to the relative clause noun (e.g., *the director that the movie...*). For passive relative clauses, the fragment contained the relative clause words up to the verb (e.g., *the director that was*

pleased...). This is the position at which the passive relatives diverge from active ones (e.g., a prompt such as *the director that* could be continued into an active or a passive) and at which the active vs. passive contrast can be best appreciated (*the director that the movie* vs. *the director that was pleased*). There were thus a total of 4 completion questionnaires, one for each of these fragments. Each questionnaire had two lists so that the animate and inanimate version of an item (e.g., *the director that...* and *the movie that...*) did not occur in the same list. Within each list, half the items contained animate heads and half inanimate heads. Each list was presented in different sequential orders and contained as many fillers as the total number of test items.

Procedure. Participants completed each questionnaire over the web and were instructed to complete each fragment provided into a sensible sentence.

Coding. Only the completed relative clause structures were analyzed, ignoring aspects of the main sentential verb phrase. Responses were coded according to two main criteria: syntactic structure and thematic roles. Among syntactic structures, we coded noun phrases, verb phrases, prepositional phrases and the order in which they occurred in the completed relative clauses. For example, for a prompt such as *the movie that* and a completion like *I saw*, we coded the relative clause structure as *np-v*. For the thematic role coding, we used standard linguistic definitions of basic thematic roles such as agents, patients, themes, experiencers, goals and locations (from Haegeman, 1991, and Belletti and Rizzi, 1988) supplemented with specific properties from Dowty (1991) because the distinction between certain roles (e.g., themes and patients) was not entirely clear from these definitions. To avoid ambiguities, the first author and a linguistics student blind to the study then assigned thematic roles to the completions according to those definitions. Although the level of agreement across coders was acceptable (about 80%), there remained ambiguities as to what was the appropriate role for a given case. We thus set further criteria to make a decision for difficult cases. For example, arguments of prepositional verbs like

talk to, would be called goal arguments, given that the action is directed to them whether or not the action affects them. Likewise, verb phrases like *is about*, *contains*, *describes*, *depicts* occurring with inanimate subjects as in *the director that the movie was about* were considered to have two theme arguments. The nouns occurring in passive constructions would be considered to bear the thematic roles of their active counterparts (as one can take the subject of a passive to bear either an experiencer or a patient role).

The definitions of the thematic roles we used were thus more precise than those given in linguistics textbooks, although they were based on the basic textbook definitions. We consider *agents* to be intentional instigators of the event denoted by the verb (Haegeman, 1990; Dowty, 1991). These often were human participants, but the label also included institutional actors such as schools and companies. For example, in *the student that the school expelled*, the school argument was considered as agent. Instigators that were themselves events such as accidents or incidents were classified as *causes*, e.g., an accident killing a person (Dowty, 1991). *Patients* were taken to be the causally affected sentient participant of events in which there was an agent (e.g., *student* in the previous example), whereas objects moving, changing possession or coming into existence (or destruction) in the event were *theme* arguments (e.g., *letter* in *writing a letter* or *destroying a letter*). This notion of theme participants includes both Jackendoff's definition of themes and Dowty's notion of incremental themes (Jackendoff, 1990; Dowty, 1991). *Goal* arguments were participants that received the theme of the event, as in *the student that the school gave the prize to*, or the argument of an intransitive prepositional verb such as *speak to* (Haegeman, 1990). *Experiencer* arguments were considered unintentionally affected sentient participants of an event or the person experiencing a state. This thematic role occurs with various verb types in different syntactic positions. Intransitive verbs such as *win* or *die* have experiencer arguments, as do experiencer-theme verbs such as *love*, *discover* and *like*, and cause-experiencer

verbs such as *annoy* and *alarm*, which are found in the experimental materials (Belletti & Rizzi, 1988; Grimshaw, 1990; Dowty, 1991). Thus experiencers differ from patients in that they undergo changes as a result of an unintentional cause. Finally, we coded as *location* nouns referring to places where the participant moving goes to or comes from and for lack of a better name, we coded adjectives like *nice* or *happy* as *adjectives*. Examples of thematic roles classifications are given in Appendices A and B.

Results

Completions at The-N-that.... Table 3 presents the completion results for animate and inanimate head prompts at all completion positions. For the *the-N-that* completions, there were two main structural choices: the relative clause was either completed into a subject relative in which the relative pronoun was followed by a verb, or an object relative clause, in which the pronoun was followed by a noun phrase but the choice of structure varied with animacy. Animate-head prompts were more often completed as subject relatives (85% of subject relatives vs. 15% of object relatives), whereas inanimate-head prompts were more often completed as object relative clauses (65% of object relatives vs. 35% of subject relatives). This choice of structure partially correlated with the choice in thematic roles. Animate heads were interpreted as agents or experiencers. The added noun within the relative clause, if any, tended to be a theme argument. In contrast, inanimate nouns were invariably interpreted as theme arguments. The most frequent second argument role for these object relative clauses, as well as the few object relative clauses with animate heads, was agent, indicating that animate agents are preferred in the subject position of an object relative clause. Overall, these results indicate a clear and early effect of animacy in sentence completion: animate and inanimate nouns are each associated with distinct thematic roles and structures.

Completions at The-N-that the.... All completions at this position had the general format of an object relative clause because the presence of the indefinite *the* after the pronoun *that* indicated that the relative clause was not a subject relative. Animate-head prompts were most likely interpreted as either patients of agent-patient verbs such as *hire* or *expel* or themes of experiencer-theme verbs such as *see* or *like* (e.g., *the employee that the manager hired/liked*), whereas inanimate heads were invariably interpreted as themes (e.g., *the movie that the director made/saw*). There were thus more alternative interpretations (more indeterminacy) in the animate-head condition. For both animate and inanimate heads, an animate noun most likely filled the subject position of the relative clause with either an agent or experiencer role. 91% of the animate-head prompts and 95% of the inanimate-head prompts had such continuations (including 13% of institutions in agent roles, e.g., *the teacher that the school hired*). The set of alternative noun and role choices available to the participants was thus perceived as quite constrained.

Completions at The-N-that-the-N. Object relative clause completions varied depending on the animacy configuration of the prompts. Inanimate-Animate configurations in the prompts, as in *the movie that the director...* were completed with a single relative clause verb in the majority of cases before moving into the main sentential verb. The roles more commonly assigned to the two prompt nouns were theme for the head-noun and agent for the relative clause subject. In contrast, animate heads prompts such as *the director that the movie* showed a variety of equally likely verb phrase structures and roles: there were passive completions such as *was written by*, in which the head noun received an agent role and the inanimate relative clause noun a theme role (*the director that the movie was written by..*), single verb completions (e.g., *distracted*, *helped*), and verb plus preposition completions (e.g., *referred to*, *belonged to*, *was for*, *was about*), in which the head noun received either a goal, theme or experiencer role and the relative clause subject

was invariably a theme of some sort. Most of the verbs used in single verb completions were ordinary agent-theme verbs used with atypical arguments. For example, verbs like *kill*, *catch* or *hit*, which normally occur with animate agents, were instead used with inanimate arguments, which were the cause of the event or the object moving between other participants (e.g., *the fish that the lure caught*, *the musician that the accident killed*, *the soldier that the grenade hit*).

Interestingly, few of these completions included cause-experiencer verbs of the type found in the original materials of Appendix A (.09 proportion of the total), indicating that only in rare occasions was a cause-experiencer interpretation formulated or expressed by the participants. Overall, the results showed marked differences across conditions: animate-inanimate relative clause configurations displayed greater indeterminacy with respect to the semantic roles, verb phrase structures and type of verb (event) to follow, whereas inanimate-animate configurations received mostly the same syntactic and semantic analysis of the original materials.

Passive completions at position The-N-that-was-Ved. These completions are shown in Table 4. Completions in passive relative clauses were fairly similar across animacy conditions. In both cases, about 70% of the passive prompts were continued with a main sentential verb and thus no constituent was added within the relative clause. The remaining 30% of the cases were by-phrases (with agent or theme-cause roles) and other prepositional phrases. The cause-experiencer verbs with animate heads were slightly more likely to be continued with a verb argument, rather than an optional adjunct such as location, indicating a tendency for these verbs to occur with the two arguments. Overall, these results show little indeterminacy compared with active object relative clauses at the same number of words into the relative clause (e.g., *the N that the N* and *the N that was Ved*): unlike active object relative clauses, the passive relative clause is perceived as already complete, even though the agent/cause role was missing, and the presence of the verb in the prompt has already assigned a role to the head noun.

Discussion

The results are consistent with the hypothesis that animacy plays a role in determining the alternative available structures and interpretations. Inanimate nouns are associated with theme-objects in object relative clause structures, whereas animate-nouns are preferentially associated with agent and experiencer roles, not only initially at the head noun position, but also later at the subject position of the object relative clause. This is not entirely surprising as animacy is a semantic feature correlated with the roles that event participants can play in an event. More interestingly, the results also support the expectation that animate-head active object relative clauses would show more indeterminacy than inanimate-head relative clauses. Animate-head object relative clauses increasingly afford more analyses across relative clause positions than inanimate-head object relative clauses. This was particularly so for the *N-that-the-N position*, the position at which the relative clause verb would be encountered during reading. The presence of an inanimate relative clause subject in active object relative clauses was followed by a myriad of interpretations, suggesting indeterminacy and difficulty in coming up with one definite interpretation. At this critical position, the animate-head object relative clauses also clearly differed from passives in the number of alternative interpretations entertained. Unlike these object relatives, passive relatives mostly received one kind of interpretation and were assigned one type of structure.

The completion preferences revealed by this study reflect what participants are likely to say in using these structures, and thus, they reflect distributional patterns in speakers' linguistic experience. Mak et al (2002) have reported that in German and Dutch corpora, subject relative clauses typically occur with animate head nouns and inanimate objects within the relative clause (e.g., *the student that recited the poem*), whereas object relative clauses almost exclusively occur with inanimate head nouns and animate relative clause subjects (e.g., *the book I read*). Similar

findings have been reported for English (Roland, Dick, and Elman, in press). These distributional patterns are consistent with the completion preferences we found for animate and inanimate head nouns. In our completions, object relative clauses tend to occur with animate subjects within the relative clause.

So far we have established effects of animacy on the indeterminacy of the interpretations. However, the main prediction of the constraint satisfaction approach is that there is a direct link between online comprehension difficulty and degree of semantic indeterminacy. We address this issue next by investigating whether indeterminacy as the relative clause unfolds can account for comprehension difficulty throughout the relative clause. Notice that mechanisms operating in an off-line sentence completion task need not occur in reading comprehension. Traxler et al. (2002) has indirectly argued for such a possibility (following findings by Ferreira et al., 2001) by suggesting that readers sometimes ignore syntactic cues such as the determiner *the* after the pronoun *that* in object relative clauses if comprehenders are strongly committed to the initial analysis and if the cues indicate that a complex structure is required. To investigate the relation between interpretation preferences and reading times, we conducted a reading comprehension study, to which we turn next.

Study 2: Reading comprehension

In this study, we examined the reading time patterns in comprehending relative clauses of different animacy configurations to evaluate our hypothesis that comprehension difficulty emerges from competition among alternative interpretations, and consequently, from the degree of semantic and syntactic indeterminacy as the relative clause unfolds. We used the materials from Appendix A, which were taken from Traxler et al. 2002 (see example in Table 1). As in study 1, we compared animate vs. inanimate-head object relative clauses, and these were in turn compared to their passive counterparts headed by the same noun. Our hypothesis predicts that

the greater indeterminacy associated with active relatives in the completions should cause more comprehension difficulty for active compared to passives, particularly at the position corresponding to the active relative clause verb, where indeterminacy is greatest for actives and is reduced for passives. Moreover, the greater indeterminacy of animate-head active clauses compared to inanimate-head ones should engender greater comprehension difficulty in the former condition. Specifically, the points of greater indeterminacy and competition with the incoming information in active structures should be particularly difficult. These points are the subject relative clause position and the relative clause verb. In our materials, the animate-head relative clauses co-occur with inanimate relative clause subjects (e.g., *the director that the movie ...*), a pattern inconsistent with the overwhelming tendency in the completions to use animate nouns in the relative clause subject position. A similar inconsistency with the completions is found at the relative clause verb position in animate-heads relative clauses (e.g., *the director that that movie pleased...*): these clauses in the experimental materials mostly occurred with cause-experiencer verbs such as *please, amuse, amaze, annoy*, in which something or someone causes a change of psychological state on a human participant - the experiencer - (Belletti & Rizzi, 1988; Grimshaw, 1990; Levin & Rappaport, 1986; Cupples, 2002). Yet in Study 1, the most common completions were agent-theme verbs. This contrasts with inanimate-heads relatives, which received continuations consistent with those of the materials (animate relative clause subjects and ordinary agent-theme verbs).

Method

Participants. Experiment participants were 63 undergraduate students at University of Wisconsin-Madison. They were native English speakers and received credit for participating in the experiment.

Materials. There were 28 items with four versions each (see appendix). The verbs and the relative clause nouns were matched for frequency and word length across conditions. The mean log frequencies for agent-theme verbs and cause-experiencer verbs were 4.13 and 3.97 respectively ($t < 1$). The mean number of characters per word for these verb types was 5.90 and 6.07 respectively ($t < 1$). The mean log frequency for animate nouns such as *director* was 4.19, and for inanimate nouns such as *movie* was 4.26 ($t < 1$), although the number of character per word differed significantly, with animate nouns being longer (7.10 vs. 5.72). The verb phrase of the main clause contained an auxiliary verb such as *has*, *had* or *was*, which was the same across conditions.

Design. There were four conditions exemplified in Table 1. Two factors were manipulated, the voice of the relative clause (active object relative or passive relative) and the animacy of the nouns (Animate-head, Inanimate head).

Procedure. A word-by-word moving-window self-paced reading paradigm was used. Reading times to each word were recorded. Participants read the stimulus sentences and immediately after answered a comprehension question about the content of the sentence. For the stimulus sentences, the questions refer to the content of the relative clause about 74% of the time. The 28 stimulus items were intermixed with 81 filler sentences (containing regular active sentences and structures with complement sentences). Each participant saw just one of the four versions of each of the 28-item sets, but across all experimental trials, participants were exposed to all four conditions. Four lists had been prepared with no more than one version of each 28 items represented in each list (thus, each experimental condition was represented seven times within each list). Sixteen different subjects saw each of the lists, except for one list that was seen by fifteen participants. Residual reading times were used as dependent variable. These were obtained from each subject's complete set of data by regressing word length onto that subject's

reading times per word and then subtracting the expected reading time from the actual reading time. Residual reading times higher than 2.5 standard deviation of the mean per word position were replaced by the cutoff value within this range. Because the active and passive relative clauses differed in number of words, we excluded from the analyses two function words in the passive relative clauses (*by the*) so that the main verb of the sentence could be aligned across all conditions.

Results

Comprehension responses. Across participants, the overall mean proportion of correct responses to all comprehension questions (including fillers) was .77. Table 5 shows means and standard deviation of the proportion of correct responses for each experimental condition. A repeated measures ANOVA with Relative Clause Voice (active vs. passive) and Animacy (animate-head vs. inanimate head) as factors and the proportion of correct responses as dependent variable revealed a main effect of Relative Clause Voice ($F(1,62)=9.87$, $MSE: .24$, $p = .003$; $F(1,27)=7.36$, $MSE: .10$, $p = .01$; $\text{minF}'(1,66)=4.21$, $p = .04$), a main effect of Animacy ($F(1,62)=11.4$, $MSE: .19$, $p = .001$; $F(1,27)=5.44$, $MSE: .23$, $p = .03$; $\text{minF}'(1,54)=3.68$, $p = .06$) and an interaction ($F(1,62)=9.7$, $MSE: .19$, $p = .003$; $F(1,27)=6.49$, $MSE: .10$, $p = .02$; $\text{minF}'(1,63)=3.88$, $p = .05$). Overall, active object relatives were more difficult than passives, and animate-head object relatives were more difficult than inanimate-head object relatives. Planned contrasts indicated that active object relatives with animate heads were more difficult than any other condition, particularly, object relatives with inanimate heads in both item and participant analyses. In the item analysis, the mean difference between the proportion of correct responses for animate-head and inanimate-head object relatives ($=.15$) had a confidence interval of $\pm .11$ ($t(27) = 2.76$, $p = .01$ two tailed). In the participant analysis, the ($=.11$) mean difference had a confidence interval of $\pm .05$ ($t(62)=4.55$, $p < .0001$, two tailed). It was the relatively low

proportion correct for object relatives with animate heads that caused the overall mean proportion across both active conditions to differ from the passive versions.

Reading times. For the analysis of reading times we aligned passive and active relative clauses as shown in Figure 1. This figure presents residual reading times for all word positions in the relative clause up to the last word of the main verb region. The results of repeated measures ANOVAs with residual reading times as dependent variable and with Relative Clause Voice (active vs. passive) and Animacy (animate head vs. inanimate heads) as factors at different regions are reported in Table 6. The first word position showing significant differences within the relative clause was the position corresponding to the un-integrated relative clause subject in object relatives, for example, *movie* in *the director that the movie pleased*. This position was aligned with the relative clause verb in the passive versions. Analyses of reading times at this position revealed a main effect of Animacy but a marginal effect of Voice. The animacy effect was carried by the object relative conditions: the animate-head object relatives were more difficult than the inanimate-head object relatives. The mean difference between these conditions was 18ms and the confidence interval for this difference was ± 12 ms ($t(27) = 3.04$; $p = .005$), whereas for the participants analysis, the mean difference and confidence interval were 19ms and ± 13 ms ($t(62) = 2.70$; $p = .005$). Planned contrasts across voice conditions, revealed that animate-head object relatives were also more difficult than inanimate-head passives (by participants analysis: mean difference = 25ms, confidence interval = ± 13 ms; $t(62) = 4.28$; $p < .0001$; by items analysis: mean difference = 24ms; confidence interval = ± 17 ms; $t(27) = 2.9$; $p = .008$) and more difficult than animate-head passives, although only marginally in the participant analysis (by participants analysis: mean difference = 14ms; confidence interval = ± 14 ; $t(62) = 1.74$; $p = .08$; by item analysis: mean difference = 17ms; confidence interval = ± 16 ms; $t(27) = 2.17$; $p = .04$) $p = .$ Inanimate-head object relative clauses did not differ from passive relative clauses.

At the object relative verb position, which was aligned with nouns in the passive relatives, there was a main effect of Voice, a main effect of Animacy and no interaction. In planned comparisons, both animate-head and inanimate-head object relatives had longer reading times than any passive condition (animate-head object relatives vs. animate-head passive: mean difference= 38ms, confidence interval= ± 21 ms., $p < .001$ for by items and by participants analyses; animate-head object relative clause vs. inanimate-head passive: mean difference: 53ms, confidence interval= ± 16 ms., $p < .0001$ in both comparisons; inanimate-head object relatives vs. animate-head passive: mean difference= 21ms., confidence interval= ± 14 ms.; $p < .006$ for both comparisons; inanimate-head object relatives vs. inanimate-head passive: mean difference= 38ms., confidence interval= ± 14 ms., $p < .0001$ in both comparisons). The two animacy conditions in object relatives only differed in the item analyses (mean difference: 18ms., confidence interval= ± 18 ms., $t(27) = 1.99$; $p = .06$). Passives relatives also differed across the animacy conditions, e.g., *movie* vs. *director*, (mean difference=17ms., confidence interval= 15ms., $p < .03$ for both by items and by participants analyses).

Next, we performed analyses in the main verb region. The region contained the word positions corresponding to the main auxiliary verb of the sentence, the following *-ed* verb and the next word. A repeated measures ($2 \times 2 \times 3$) ANOVA was performed on this region with residual reading times as dependent variable and with Relative Clause Voice (active vs. passive) and Animacy (animate vs. inanimate heads) as crossed factors and word positions as nested factor (main auxiliary verb, main verb, and next word). Results are reported in Table 6. The main effect of word position was due to the fact that across word positions, the main auxiliary verb showed the largest reading times in all conditions. Likewise, the interaction of word position and sentence type was due to the fact that although overall actives took longer than passives, the magnitude of the effect varied with word position.

Analysis of variance at the main auxiliary verb position revealed effects similar to those in the entire region. There was a main effect of Relative Clause Voice, ($F(1,62)=122$, $MSE:140213$, $p < .0001$; $F(1,28)=121$, $MSE:65384$, $p < .0001$, $\min F'(1,77)=60.7$, $p < .0001$) a main effect of Animacy ($F(1,62)=7.96$, $MSE:14498$, $p = .006$; $F(1,28)=11.7$, $MSE:9912$, $p = .002$, $\min F'(1,87)=4.78$, $p = .03$) and an interaction ($F(1,62)=5.29$, $MSE:8483$, $p = .02$; $F(1,28)=6.73$, $MSE:4909$, $p = .01$, $\min F'(1,84)=2.96$, $p = .09$). As before, active object relatives were more difficult than passives, and animate-head object relatives were more difficult than inanimate-head object relatives (mean difference= 27ms., confidence interval= 17ms., $p < .001$ for by items and the by participants analyses). Contrasts across voice conditions revealed that both animate and inanimate-head object relatives were more difficult in this position than passive relatives. The mean differences for these comparisons were larger than 32ms. and the confidence intervals were all larger than 10ms. (17ms. and 15ms. for animate-head object relatives vs. animate-head passives in the by items and by participants analyses respectively; 10ms. and 12ms. for inanimate-head object relatives vs. inanimate-head passives in the by items and by participants analyses respectively, $p < .0001$ for all comparisons). Passive relatives did not differ from each other.

Analyses at the main *-ed* verb of the main clause revealed the same pattern of results as the auxiliary word position. There was a main effect of Relative Clause Voice ($F(1,62)=24.8$, $MSE:41191$, $p < .0001$; $F(1,28)=40.89$, $MSE:20620$, $p < .0001$, $\min F'(1,88)=15.4$, $p < .0003$), a main effect of Animacy ($F(1,62)=14.8$, $MSE:28549$, $p = .0003$; $F(1,28)=17.35$, $MSE:14113$, $p = .0003$, $\min F'(1,88)=7.98$, $p = .006$) and an interaction only in the item analysis ($F(1,28)=6.59$, $MSE:4247$, $p = .01$). The same pattern was also observed for the next word (main effect of Voice: $F(1,62)=21$, $MSE:21278$, $p < .0001$; $F(1,28)=18.65$, $MSE:7739$, $p = .0002$, $\min F'(1,74)=9.87$, $p = .003$); main effect of Animacy: $F(1,62)=16$, $MSE:15651$, $p =$

.0001; $F(1,28)=12$, MSE: 6880, $p = .0016$, $\min F'(1,68)=6.85$, $p = .01$). At the main verb position, planned comparisons revealed that the animate-head object relatives were more difficult than any passive condition (mean differences were larger than 35ms., with a confidence interval of ± 20 ms. and $p < .001$ for all comparisons). Animate-head object relatives were also more difficult than inanimate-head object relatives (mean difference= 31ms., confidence interval= ± 18 ms., $p < .002$ for by items and by participants analyses). However, passive relatives did not differ themselves and inanimate-head object relatives only differed from inanimate-head passives (mean difference= 16ms., confidence interval= 9ms., $p < .003$ for by items and by participants analyses). A similar pattern for results was obtained for planned comparisons at the next word position.

In sum, using both proportion of correct responses and reading times measures, active object relatives with animate heads were more difficult than any other condition, most notably, than active object relatives with inanimate heads. This difficulty starts to be observable early within the relative clause at the embedded noun position and becomes larger at the main auxiliary verb of the sentence, spilling over the next two words. The passive versions of these relatives were much easier to process in comparison, with little differences between the animacy conditions.

Discussion

As expected on the basis of indeterminacy, these results indicate that animate-head object relatives are more difficult to process than inanimate-head object relatives, and that the points of high indeterminacy or competition as established by the completion results correspond to points of greater difficulty. Unlike in previous studies, which either grouped together words within the relative clause (Traxler et al (2002) or took the relative clause verb to be the first critical word (King & Just, 1991; Gordon et al, 2001; Grodner and Gibson, 2005), the analyses here show that the difference across animacy conditions is observable at the earliest possible position within the

relative clause, suggesting an immediate effect of animacy. The animacy of the relative clause subject in the animate-head condition (e.g., at *the director that the ... movie*) must have caused competition between alternatives, as the previously entertained interpretations (e.g., at *the director that the...*) were inconsistent with the interpretations suggested by the current input. In contrast, inanimate-head relatives show little evidence of difficulty at this position, as they are no more difficult to process than the passives. The relative clause verb in the animate-head condition also show the most difficulty, in agreement with the greater indeterminacy found at the *N-that-the-N* position in the completions. The competition at this position continues into subsequent words, which themselves introduce new thematic roles information.

The results also indicate a large difference across active and passive relatives. In particular, inanimate-head object relatives, which have been previously shown to be no more difficult than typical active subject relatives, differed from passive relatives at various positions, most notably at the relative clause verb (as expected) and the main sentence verb, a position thought to reflect semantic integration across the relative clause words (King & Just, 1991, Gibson, 1998).

Although one should be careful in interpreting these data due to the fact that active and passives are not exactly matched in number of words, this finding is consistent with indeterminacy in that after the *by*-phrase of the passive relatives has been read, there is little indeterminacy, as no other constituent is expected before the main verb, and the interpretation of the event described by the relative clause has been completed. This contrasts with the active relatives in that the event described by the relative clause (with specific participant roles) is still being processed.

Up to now, we have pointed out correspondences between comprehension difficulty and indeterminacy only informally. To investigate the relation between interpretation preferences and reading times more precisely, we performed a series of regression analyses in a systematic fashion. In doing so, we tested more specific predictions of the indeterminacy account.

Study 3: Predicting Reading times

In this study, we investigate whether the interpretation preferences of the gated sentence completions predict reading times at each word position within the relative clause. We tested the hypothesis that semantic and syntactic indeterminacy contribute to comprehension difficulty due to competition among likely alternatives. To evaluate this possibility, we regressed the proportion of interpretations from Study 1 at different relative clause positions onto the reading time of subsequent positions from Study 2. The likelihood of an analysis can be an index of indeterminacy because the more likely an interpretation, the less competition from alternatives and the less difficult to activate. Inversely, the less likely an interpretation, the more difficult to activate and the more the competition with other frequent alternatives. We also regressed onto the reading times the number of alternative interpretations entertained at each point and the number of verbs provided in the completion study. These measures provide an additional index of indeterminacy: the more interpretations or verbs were provided in the completions for a given item, the more the alternative interpretations available to compete.

We predicted the following. First, the reading times at a given position would be partially explained by the completions up to this position. The degree to which speakers did or did not entertain analyses similar to the upcoming one (including the animacy of the upcoming noun) should correlate with the ease or difficulty of activating the upcoming structure. Similarly, at high indeterminacy points, the activation of more competing analyses should correlate with more comprehension difficulty. This means that the frequency distribution of preferred interpretations at *the-N-that...* and *the-N-that-the...* should predict the reading times at *the-N-that-the-N* and the interpretations at and *the-N-that-the-N* should predict the reading times of the subsequent verbs. Second, we reasoned that because the same syntactic configuration can receive different thematic role interpretations, the thematic completions have the potential to account for more variance

than the syntactic alternatives. In particular, the subject vs. object syntactic alternatives at *the-N-that* position, which was stressed by Traxler et al. (2002) as a significant source of ambiguity, is predicted to be less critical than broader semantic indeterminacy. Thus if comprehenders are sensitive to semantic indeterminacy in early relative clause positions, then the alternative thematic roles and animacy configurations at the *the-N-that* and *the-N-that-the* should be better predictors of reading times at the relative clause subject position than the corresponding syntactic alternatives. Likewise, because of the argument-taking function of verbs and their critical role in resolving thematic indeterminacy, competition between alternative interpretations at verb positions (relative clause verb and main verb phrase) should be better predicted by semantic measures than structural ones at any position. Finally, the distributions of active and passive reading times taken together should be predicted by their varying degrees of semantic indeterminacy at all relevant positions.

Procedure. For each sentence completion position, we used the coded completions of Study 1 to compute several predictors that were then regressed onto the reading times of Study 2 at each relevant position. We first calculated three different measures, one based on syntactic structure alone, the second on thematic roles, and the third on the animacy of the completed noun (for the case of *the-N-that-the* position). Specifically, for each completion position and each item, we computed the proportion of completions that were consistent with the ultimately “correct” analysis of the relative clause of this item in Study 2, i.e., completions consistent with the item’s upcoming relative clause structure and interpretation in Study 2 (cf. Appendix A). For brevity, we will refer to these proportions as the proportion of “correct” analyses. For the structural count, we computed the proportion of cases in which each item (e.g., *the director that...*, *the director that the...*, *the director that the movie...*) was continued into a structure like *the-N-verb* (*the director that the movie pleased*). For the first completion position, a “correct”

relative clause structure was *np-vp*, regardless of the complexity of the verb phrase because this captures the structure of the upcoming words after *the-N-that* position. In later positions, the internal verb phrase structure was taken into account and *np-v* was considered “correct”, because the activation of structures such as passives (e.g., *np-aux-v-p* in *the director that the movie was written by*) or prepositional structures (*np-v-p* in *the director that the movie was based on*) are likely to compete when the relative clause verb is encountered. The same procedure was used for inanimate-head items. In addition, to predict reading times in passive relative clauses, we used the completions at *the-N-that* that had passive-like structure (*(auxiliary)-v-adjective*, *(auxiliary)-v-adv*, *(auxiliary)-v-pp* to compute the proportion of “correct” passive responses. Appendix B provides an example of how different predictors were obtained from the structural and thematic coding.

For the thematic role count, “correct” proportions included thematic roles that were close in meaning to the ultimately “correct” ones indicated in appendix A. Because the cause-experiencer argument relations found in the materials were not used at early relative clause positions such as *the-N-that* and *the-N-that-the*, the “correct” role for the head noun in these structures was taken to be experiencer, patient or goal followed by an agent or cause noun (the semantically closest configuration to the correct experiencer-cause one). For *the-N-that-the-N* position, the “correct” roles for these nouns were taken to be experiencer-cause, experiencer-theme respectively. For the inanimate-head relatives, the “correct” role for the head noun was theme, followed by an agent or experiencer role, as in the stimulus materials. For passive relatives, we used the coded roles at the *N-that* that were like those in passives to compute the proportion of “correct” roles. For animate-head passive, experiencer and patient role of the noun head followed by agents, causes or adjectives were considered “correct” roles as these are similar to the experiencer-cause roles of the passives in the materials. For inanimate-head passives, a theme role assigned to the head

noun followed by an agent or experiencer role was considered “correct” (see examples in Appendix B).

To obtain an additional index of semantic indeterminacy, we also computed two non-proportional measures for each item. One was the number of alternative thematic role interpretations provided in the completions. For example, if an item at the *N-that-the-N* position received a *theme-agent* interpretation in half the cases, and a *theme-experiencer* interpretation in the remaining cases, we coded two alternative interpretations for that item. The other measure was the number of verbs used at *the-N-that....* and *the-N-that-the-N* positions for a particular item. For example, prompts such as *the director that the movie....* tended to be completed with certain common verbs and sometimes participants agree in which verb this would be (e.g., *was about, was written by...*). The number of verb phrase continuations is thus an index of how many alternative interpretations were entertained in the completions across participants. The rationale for using this latter measure was that it could allow us to compare indeterminacy of verb predications (uncertainty of one or another specific verb) to the more conceptually based indeterminacy of the relation between event participants, which is captured by the thematic role measure. Because animate entities can participate in many more events than inanimate entities, they generally co-occur with more verbal predicates (Bock, Loebell, and Morey, 1990). This fact could potentially influence indeterminacy, as many more verbs could be entertained for animate-head relatives compared to inanimate-head ones.

Results

Active object relative clauses Table 7 shows the correlation matrix resulting from regressing the number of alternative interpretations and verbal predicates, and the structural, thematic and animacy completion proportions at each relative clause position onto the reading time of subsequent sentential positions. All *p* values are smaller than .05 and different *p* values are

indicated with stars. Bonferroni correction for multiple individual regressions on the same reading time variable renders the significant value between .01 to .006 depending on the number of analyses. Nevertheless, because this correction is fairly conservative, we have included values at $p = .02$ in the table for information. Non-significant or non-applicable correlations have been left blank. Examples of scatter plots with regressions at different word positions are shown in Figure 2.

The reading times of the first content word within the relative clause (e.g., *movie* in *the director that the movie*) was predicted by various completion measures: the proportion of “correct” animacy completions and the proportions of “correct” syntactic and thematic interpretations entertained at *the-N-that* and *the-N-that-the* positions, indicating that the more “correct” analyses were entertained, the less the reading time difficulty. (One outlier was removed from this dependent variable using case diagnostics procedures described in Field, 2005). To test whether semantic predictors significantly increased the amount of variance accounted for by syntactically based ones, we performed a multiple regression in a step-wise fashion and examined the change statistics. A regression model with only the structurally “correct” analyses at *the-N-that* accounted for 38% of the variance, whereas the addition of the number of alternative interpretations at *the-N-that-the* into the regression model increased the variance to 56%, a significant increase ($F(1,52)=13.25$, $p = .001$). The proportion of “correct” animacy completions entertained at *the-N-that-the* also had this effect: the addition of this predictor to a model containing the proportion of structurally “correct” analyses at *the-N-that* increased the to 47% ($F(1,52)=4.93$, $p = .03$). This indicated that the animacy and thematic roles entertained right before the relative clause subject was encountered were a better predictor of difficulty at this position. By contrast, the addition of the structural predictor to a model already containing these semantic predictors resulted in no significant increase ($F(1,52)=.9$).

Interestingly, the proportion of correct thematic interpretations at the *the-N-that-the* position was also a better predictor than the proportion of “correct” animacy completions at this position. A model containing the animacy predictor accounted for 12% of the variance, whereas the addition of the thematic role predictor increased the variance accounted for to 21%, a significant increase ($F(1,53)=6.36, p=.01$). In contrast, animacy does not increase the variance accounted for when the thematic role predictor is already in the model. This indicates that the roles that the nouns played in the event have more weight than simply noun animacy in predicting comprehension difficulty. The more their semantic roles are like those ultimately correct, the faster the reading time.

Table 7 shows that a similar prevalence of semantic analyses was observed at the relative clause verb position. Reading times at this position were positively though weakly correlated with the number of alternative interpretations at *the-N-that-the-N*, indicating that some thematic role information was processed at this position. This weak trend continues into later relative clause positions. More alternative interpretations entertained entail more comprehension difficulty. Interestingly, the relative clause verb position was also positively correlated with the number of alternative predicates used in the completions, indicating that the more potential verbs (events) were entertained for a given prompt at *the-N-that-the-N* position, the more the reading difficulty.

At the main auxiliary verb, the number of thematic alternatives and the thematic role count at position *the-N-that-the-N* together accounted for 18% of the variance. The proportion of “correct” thematic roles at the *the-N-that-the-N* position was a better predictor than the structural count at *the-N-that*, indicating that entertaining interpretations similar to the ultimately correct one correlates with less comprehension difficulty. A regression model with only the structural “correct” analyses at *the-N-that* accounted for 9% of the variance, whereas the addition to the

model of the number of alternative interpretations increased the variance to 16%, a significant increase ($F(1,53)=4.27$, $p = .04$). In contrast, the addition of the structural predictor to a model containing the semantic predictor resulted in no significant increase ($F(1,53)=.06$).

Likewise, the main verb position, the thematic role count at position *the-N-that-the-N* was the best predictor. A regression model with only the structural “correct” analyses at *the-N-that* accounted for 26% of the variance, whereas the addition to the model of the thematic analyses at *N-that-the-N* increased the variance to 37%, a significant increase ($F(1,53)=9.77$, $p = .003$). By contrast, the addition of the structural predictor to a model containing the thematic predictor resulted in no significant increase ($F(1,53)=1.31$). The reading times of the main verb were the ones best accounted for by the thematic completion measure, indicating that competition between alternative roles for the relative clause nouns took place at this position to a greater degree, a process that had started at the relative clause verb position.

Overall the results indicate that semantic preferences with its concurrent indeterminacy had an immediate effect on reading times throughout the relative clause, with larger effects at the main verb phrase. This suggests that although competition between alternatives takes place throughout the relative clause, competition is stronger at verb positions, at which the “correct” interpretation can start being activated, and is not resolved until the main verb phrase.

Active and Passive Relative Clauses. When active and passive relative clauses were analyzed together in the same regression, variation in reading times across conditions was predicted by thematic but not structural preferences at position *the-N-that* resembling the ultimately correct roles of active object relative clauses or passive relative clauses respectively. The “correct” active and passive thematic preferences at *the-N-that* position accounted for significant variance at three word positions: (1) at position 5 in Figure 1 ($r = -.25$, $p = .008$), where the relative clause verb in passives and the subject noun in object relative clauses were aligned; (2) at the main

auxiliary verb ($r = -.22$, $p = .02$) and (3) at the main verb ($r = -.39$, $p = .0001$). These results indicate that the activation of the appropriate or inappropriate roles in all four relative clause versions explains reading difficulty, and that the active and passive relative clause differ in indeterminacy: the proportion of ultimately correct thematic interpretations in the completions predicted longer reading times in active and passive structures. The less likely an interpretation, the longer the reading times and vice versa.

Discussion

The results of the regression analyses largely confirm the predictions of the constraint-based approach. The degree to which the completion interpretations were like those ultimately correct at a given relative clause position predicted their comprehension difficulty at subsequent positions and across relative clause types. This was particularly so for points of high indeterminacy such as the subject relative clause and the main verb phrase. The only position at which structure and some semantic measures did not predict reading times was the relative clause verb. This position was only predicted by the number of interpretations entertained and by the number of verb predicates provided in the completions. This may indicate that the semantic representation being entertained at this position were input-driven schematic representations rather than detailed ones. This is consistent with the great difficulty and the strong correlations obtained at subsequent positions, where semantic roles were being computed.

In addition, semantic measures were stronger predictors than the structural measures in both passive and active relative clauses at all positions, indicating that thematic indeterminacy had an immediate effect on reading times. Because the semantic indeterminacy captures more fine-grained distinctions about event roles (as there may be several semantic analyses for the same structure), the coarse structure alone misses this type of information. This was particularly noticeable at verb positions, at which the completion showed greater indeterminacy and the least

frequent ultimately “correct” interpretations for the animate-head condition. The activation of the correct interpretation was thus slower and competition stronger at verb positions computing arguments’ roles, spanning several word positions. The thematic role measure was also a stronger predictor than animacy, indicating that although thematic roles are related to the animacy of the nouns, they capture other properties such as the relationship between event participants that are not included in noun animacy. Overall, the degree of competition between alternative interpretations substantially contributes to the degree of comprehension difficulty, thus explaining comprehension differences across relative clause types.

General Discussion

This article investigated the claim that a major source of difficulty in object relative clause comprehension is indeterminacy in interpretation of the sentence as it unfolds over time. This approach contrasts with alternative views in which the syntactic complexity of the clauses, not their ambiguity, was the source of comprehension difficulty. A series of studies investigated object relative processing from a constraint-based ambiguity resolution perspective. Study 1 used the gated sentence completion task to assess which alternative interpretations are dominant as the sentence unfolds, Study 2 compared reading times in object relatives to an unambiguous control condition, passive relatives; and Study 3 related completion data and reading data. The results showed longer reading times for object relatives compared to unambiguous passive relatives. These differences were well correlated with the availability of alternative interpretations, as revealed by the completion data. Our results thus provide evidence for the hypothesis that semantic and syntactic indeterminacy in relative clause interpretation is a source of comprehension difficulty in these structures. These results argue for an approach to relative clause comprehension within a constraint-based approach to ambiguity resolution, in which comprehension difficulty is a function of the amount of indeterminacy in the sentence at various

points in time. This indeterminacy is itself a function of the extent to which lexical and other constraints converge to promote a single, ultimately correct interpretation.

This approach, aligning relative clause processing with well-attested examples of constraint-based ambiguity resolution, effectively seeks to place the two major subfields of sentence processing research—syntactic ambiguity resolution on the one hand, and complex relative clause interpretation on the other—within a single approach of constraint based comprehension processes. This approach leads naturally to questions concerning the effects of this reconceptualization on the currently largely distinct fields of ambiguity resolution and relative clause processing, which we address below.

Relationship to constraint-based ambiguity resolution accounts

A key concept of constraint based sentence processing is that various types of linguistic information (words, phrases, intonation contours, etc.) provide probabilistic constraints, that is, information that affects the likelihood of the interpretation of the current and upcoming linguistic signal. A number of studies within ambiguity resolution research have argued that comprehension difficulty varies as a function of the extent to which constraints converge or conflict in support of various sentence interpretations (Tanenhaus & Trueswell, 1995; Mc Rae et al, 1998). If object relative interpretation is an example of such processes, then it should be possible to identify a number of probabilistic constraints that should modulate the difficulty of interpretation of these structures. The present results and those of other recent studies point to several potentially important constraints. First, this study and others (Mak et al. 2002; Traxler et al. 2002) have shown the importance of head noun animacy in relative clause interpretation, in that object relatives with inanimate head nouns are easily interpreted, but object relatives with animate head nouns are not. Within the constraint-based account, noun animacy provides probabilistic information modulating the relative likelihood of alternative interpretations.

Animate nouns most often receive agentive or experiencer interpretations, whereas inanimate nouns are commonly the theme-object of an action. This is not only a property of the linguistic input in speakers' experience but also a property of the speakers' world, which is dominated by human activity. It is thus not surprising that animate nouns tend to be sentential subjects and that an agentive interpretation is initially adopted when a sentence-initial noun is encountered. For object relatives, indeterminacy later arises when the unfolding evidence conflicts with this agentive interpretation for the animate noun. In contrast, a causal (agentive-like) event role rarely co-occurs with inanimate nouns (either in linguistic experience or in the world). Inanimate nouns in clause-initial position (of main or relative clauses) are thus less likely to receive such an interpretation. The comprehender's commitment to non-causal interpretations of inanimate nouns then engenders extra processing cost on those relatively rare occasions when a causal interpretation of inanimate nouns is ultimately correct.

This interpretation of animacy effects as modulating interpretations in relative clause processing is consistent with work in syntactic ambiguity resolution, in which noun animacy is well known as an important probabilistic constraint in interpreting syntactically ambiguous relative clauses (Trueswell et al., 1994; McRae et al. 1998). In these cases, as with the object relatives, noun animacy affects the probability of the noun being the agent or patient/theme of some upcoming verb. Our results in addition indicate that although animacy influences the type of semantic roles that event participants play, the relationship between participants in the event described is a stronger constraint than noun animacy alone.

Another potential constraint concerns the noun vs. pronoun status of embedded subjects in object relatives, as in the contrast between, *The reporter that the senator attacked...* and *The reporter that he attacked...* Several researchers have found that object relatives are substantially easier to comprehend when the embedded subject is a pronoun (Gordon et al., 2001; 2004;

Warren & Gibson, 2002), but the explanation for this result is a subject of debate. Warren & Gibson (2002) have argued that object relatives containing subject pronouns are easy to process compared to those with noun subjects because pronouns are readily available in the discourse representation, thus decreasing memory load. Gordon et al., however, have attributed the effect to varying degrees of interference between the head noun (e.g. *reporter*) and the embedded subject (*senator/he*); they hypothesized that pronouns interfere less with head nouns by virtue of being a different lexical category, resulting in reduced processing difficulty compared to the more interfering noun subject condition. Certainly nouns and pronouns have different discourse statuses, and there is abundant evidence that discourse information can affect interpretation of ambiguous sentences (e.g. Altmann and Steedman, 1988). On this view, Warren and Gibson's point about discourse status and processing difficulty could be made to fit naturally into a more constraint-based framework. However, Real and Christiansen (2007) have recently provided another explanation: They argue that object relatives with pronominal subjects are relatively easy to process in part because pronominal subjects in object relatives are highly frequent in the corpus (probably owing at least in part to their special discourse status). Our present results are also consistent with this view. In Study 1, about 82% of the object relative completions at the *the-N-that* position contained personal pronouns such as *I*, *she*, *he*, or *you*. Together with previous corpus studies of relative clause use (Jaeger & Wasow, 2005; Race & MacDonald, 2003), all these results suggest that linguistic experience with certain structural patterns influence reading comprehension difficulty. Phrased in terms of indeterminacy, the presence of a pronoun following a head noun, as in ...*the reporter (that) I*... strongly predicts an object relative clause interpretation of the string.

There are likely a number of other constraints that modulate the ease of comprehension of object relative clauses in addition to the ones that we have identified here. Such constraints are

also likely to increase the amount of variance explained so far. For example, Race and MacDonald (2003) combined corpus analyses and self-paced reading studies to investigate the effect of distributional regularities of word choices on the comprehension of object relatives. They found that readers were highly sensitive to the distribution of relative pronouns (e.g. *that*); reading times in object relatives were shorter when the sentence contained particular combinations of relative pronouns and embedded subject types that were highly frequent in object relatives, compared to conditions in which the word combinations were rare in object relatives. These results can be interpreted in terms of indeterminacy; word patterns that in the past have frequently co-occurred with object relative clause interpretations strongly favor the object relative interpretation when these word patterns are encountered in new input.

Another constraint affecting relative clause interpretation is likely to be properties of the verb in the relative clause. In related work, we have argued that the verb's lexical type can contribute additional difficulty in processing object relative clauses because of frequent patterns of thematic roles assignments and syntactic frames associated with the verbs (Gennari & MacDonald, 2003). Cause-experiencer verbs of the sort used in the materials of our studies tend to occur in passive constructions (Ferreira, 1994) and are more difficult to process when they occur in actives (Cupples, 2002). Consequently, speakers completed our prompts with relatively few cause-experiencer verbs, and comprehenders experienced difficulty in interpreting them in active object relative clauses. Indeed, the passivization rate of cause-experiencer verbs in speakers' productions accounts for a significant amount of variance in reading times at the main verb region of object relative clauses, indicating difficulty in assigning thematic roles within the current noun configuration (Gennari & MacDonald, 2003). Competition between the verbs' frequent pattern of role assignments and alternative structures activated by the unfolding input

thus very likely increased the difficulty already caused by thematic indeterminacy up to that point in the relative clause.

Probabilistic constraints contributing to the ease or difficulty of interpreting object relative clauses have often been criticized for putatively making the wrong predictions. Grodner and Gibson (2005) for example, have argued that experience-based accounts would predict comprehension difficulty at the first word position within object relatives but not at subsequent verb positions. Their critique was directed at the experience-based modeling of MacDonald and Christiansen (2002), who emphasized properties of object relative structures and did not consider the role of experience in other constraints discussed here, such as noun animacy, verb type, etc. (see Wells et al., 2006, for discussion). However, it follows from the verbs' lexical meanings and their argument-taking function that all other material in the preceding partial structure would have to be semantically composed with them. Verbs are thus the places where thematic indeterminacy and verbs' lexical patterns matter most.

Other studies have also pointed out the difficulty of determining the appropriate grain of frequency-based information, which makes experience-based approaches difficult to falsify (Gordon et al, 2004; Grodner, Gibson and Tunstall, 2002; Mitchell, Cuetos, Corley, and Brysbaert, 1995). This is undoubtedly a problem in that not every frequency count will provide useful information in a particular construction. However, an extensive literature on constraint-based models has already identified a number of factors that contribute to comprehension difficulty, including lexically driven biases such as noun animacy and thematic roles (see Trueswell et al, 1994), verb argument structures or syntactic frames (e.g., Boland et al, 1995; Trueswell, Tanenhaus & Kello, 1993), world knowledge biases such as plausibility (MacDonald 1994, Pearlmuter and Mac Donald, 1995), contextual information (e.g., Trueswell & Tanenhaus, 1991), and co-occurrence patterns (MacDonald & Christiansen, 2002; MacDonald, 1993). Some

of these studies have pointed out that indeterminacy or ambiguity of interpretation, along with that of structure, is critical in explaining comprehension (e.g., Boland et al. 1995; Tanenhaus, et al., 1989; Trueswell et al., 1994). There is thus no principled reason why factors that contribute to ambiguity resolution in "traditional" syntactic ambiguities should not make any contributions to comprehension of other constructions, as the cues to ambiguity resolution are also cues to entertain one or another interpretation of technically unambiguous constructions. This greater array of constraints may prove useful in revising assessments that expectation-based accounts (constraint-based and information-theoretic approaches) do not provide a good account of relative clause processing (Levy, 2007), in that a richer array of constraints of the sort discussed here may yield a different set of predictions for the locus of processing difficulty than has been considered in that literature to date. Of course this rich array of constraints proposed here carries with it a burden to explain in a mechanistic way how they are integrated and weighed during sentence comprehension, and that remains a challenge for constraint-based accounts.

In sum, we have identified some probabilistic constraints that modulate the difficulty of interpreting object relative clauses, and there are likely to be a number of others. It is beyond the scope of this work to address why these particular constraints obtain, that is, why object relative clauses are particularly associated with inanimate heads, pronominal subjects, certain kinds of verbs, etc. In other work we have argued that these and other distributional patterns emerge because of constraints on the language production system (Gennari & MacDonald, 2006; MacDonald, 1999; Race & MacDonald, 2003). Our point here is that whatever their ultimate origin, these constraints are robustly represented in the input, they are learned from prior experience with relative clauses and other co-varying linguistic structures (Wells, Christiansen, Race, & MacDonald, 2006; Kidd, Brandt, Lieven and Tomasello, 2007), and when a new relative clause is encountered, they guide the interpretation of this new input by modulating the

likelihood of alternative interpretations. This account emphasizes both learning from experience and the richness of that experience, with a number of probabilistic constraints contributing to the ease or difficulty of interpreting object relative clauses.

Relationship to Other Accounts of Relative Clause Processing

As described in the introduction, most accounts of object relative clause processing have assumed that object relatives are essentially unambiguous, and the accounts have viewed object relative difficulty as stemming from high memory demands in parsing these complex structures. A recent minority view is that some ambiguity is contained in the structures, and difficulty stems from misanalysis of object relatives (Traxler et al., 2002). Our approach is distinct from both of these alternatives.

The present findings challenge Traxler et al.'s (2002) two-stage parsing proposal that all relative clause heads are initially analyzed as syntactic subjects, with animacy playing a role only during reanalysis. Inanimate head object relative clauses are unlikely to be initially mis-analyzed as subject relatives: the preference for object relative structures in the earliest stage of the gated sentence completions in Study 1 and the rapid reading times in the subject relative clause position in Study 2 (as rapid as passive relatives) shows little evidence of difficulty and thus of reanalysis. In addition, the number of interpretations at *the-N-that-the* and noun animacy were better predictors of comprehension difficulty than were syntactic preferences at the relative clause subject position. Yet it can be argued that the reanalysis occurs so fast that it is undetected or that reanalysis in fact takes place at the verb positions, even though the subject relative clause noun is read fast (the reanalysis proposal is not very specific). Taken these observations together, our results indicate that semantic indeterminacy plays at least as strong a role as the initial subject-object relative ambiguity in explaining comprehension difficulty. This conclusion is consistent with eye tracking data from Mak et al. (2006), who also failed to find evidence of

reanalysis in comparing active subject and object relatives with similar animacy configurations and with neuropsychological evidence showing that even patients with syntactic impairments are sensitive to semantic constraints (Saffran and Schwartz, 1994). Thus while we agree that ambiguity resolution is central to understanding object relative interpretation, it remains an open question whether all object relatives are inevitably misanalyzed as subject relatives.

Our results also provide a way to reconceptualize memory-based approaches that explain comprehension difficulty in terms of distance between noun-verb relations (Gibson, 1998, Grodner et al., 2002; Grodner & Gibson, 2005) or similarity-based retrieval at integration points (Lewis and Vasishth, 2005; Van Dyke and Lewis, 2003; Van Dyke, 2007; Gordon et al 2001, 2004). These researchers have viewed long reading times and/or poor comprehension accuracy in object relatives as evidence of memory costs during processing, while we have interpreted these data as stemming from the activation of alternative interpretations during processing. These contrasting positions reflect different degrees of attention to lexically-specific comprehension processes. That is, the memory based accounts view comprehension difficulty as owing to basic properties of the human memory system, while our approach, and constraint-based ambiguity resolution accounts more generally, emphasize the degree to which the activation of alternative interpretations vary as a function of the particular properties of the sentence, such as noun animacy, verb meaning, thematic fit between nouns and verbs. Thus while at some level of explanation it must be the case that the architecture of human memory shapes comprehension processes, our approach offers a finer-grained account of comprehension difficulty tied to particular properties of the sentence to be comprehended. One way of conceptualizing the two general approaches is that the memory based accounts are essentially correct that the comprehender faces interference in maintaining activation of information while processing object relative clauses, but that a major source of this interference is the competition

from alternative interpretations that are partially activated as a function of constraint based ambiguity resolution processes.

On this view, semantic indeterminacy bears some resemblance to the indeterminacy in the noun-verb relationships cited in memory-based accounts (e.g., Gibson, 1998; Van Dyke and Lewis, 2003), as the difficulty of assigning thematic roles may be correlated with the difficulty of establishing a distant link or retrieving the corresponding nouns from memory when verbs are encountered. However, the focus on similarity-based retrieval or on quantitative measures of distance such as the number of intervening un-integrated nouns, largely independent of the lexical properties of the sentence to be comprehended, falls short of explaining the current animacy effects. Measures of complexity based on distance or similarity predict that object relative clauses with animate and inanimate heads such as those in our materials should be equally difficult, as the distance from the verb and the nouns' similarity stays constant across conditions. Recent memory-based research has begun to incorporate small amounts of sentence-specific information in predictions of processing complexity, such as the distinction between object relatives with pronoun and noun subjects (Gordon et al., 2001; 2004; Warren & Gibson, 2002). Our approach goes well beyond this limited attention to lexical items and in addition suggests that lexical properties of words occurring in certain syntactic configurations affect the activation of alternative interpretations during comprehension.

In this respect, the semantic indeterminacy approach has implications for the classical memory-based complexity contrast between subject and object relatives. In subject relative clauses, like in our passive clauses, the relative clause verb is encountered earlier than in object relative clauses. The early semantic composition of the verb with its subject argument thus reduces semantic indeterminacy within the relative clause and triggers the activation of potential roles for the upcoming object, if the lexical meaning of the verb so requires (Boland et al, 1995;

Tanenhaus, et al., 1989). In object relative clauses, in contrast, the head noun and the relative clause subject activate competing roles and ways in which they can be related into an event, engendering more indeterminacy and difficulty. On this view, different degrees of indeterminacy may be responsible for some amount of the increased difficulty of object relatives compared with subject relatives. However, as we noted in the introduction, the sentences also differ in meaning, which may also contribute to variations in processing difficulty.

Finally, semantic indeterminacy can also shed some light on results that have been recently reported in Mak et al, (2006). The studies reported in this work examined several animacy configurations, including inanimate-inanimate combinations. These authors attribute comprehension difficulty to topichood cues such as subject function and animacy. Although this account is consistent with a constraint-based approach, thematic role indeterminacy suggests the possibility that inanimate nouns (some of which refer to events such as *accident* or *leakage*) in certain positions may activate different thematic roles as a function of their meanings, thus possibly explaining some of the preferences found in Mak et al., (2006). For example, event-referring head nouns in relative clauses may have a tendency to be interpreted as causes, thus, engendering a subject relative preference. Such a possibility remains to be investigated.

Author Notes

This research was supported by NIMH grant P50 MH644445, NICHD grant R01 HD047425 the Wisconsin Alumni Research Fund. We thank Scott Spilger for his efforts coding the completion data in Study 1 and anonymous reviewers for helpful comments on an earlier version of this work.

References

- Altmann, G., and Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30(3), 191-238.
- Altmann, G. T. M., and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Baird, R., and Koslick, J. D. (1974). Recall of grammatical relations within clause-containing sentences. *Journal of Psycholinguistic Research* 3(2), 165-171.
- Belletti, A., and Rizzi, L. (1988). Psych-verbs and θ -theory. *Natural Language and Linguistic Theory*, 6(3), 291-352.
- Bock, J. K., Loebell, H., and Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99, 150–171.
- Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, 12(4), 423-484.
- Chierchia, G. (1989). Structured meanings, thematic roles and control. In G. Chierchia, B. H. Partee and R. Turner (Eds.), *Properties, Types and Meaning, II* (pp. 131-166). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Chierchia, G., and McConnell-Ginet, S. (1990). *Meaning and grammar: An introduction to Semantics*. Cambridge, MA: MIT Press.
- Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., and Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, 34(6), 774-806.
- Caplan, D., and Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22(1), 77-126.

- Clifton, C., and Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. N. Carlson and M. K. Tanenhaus (Eds.), *Linguistic Structure in Language Processing* (pp. 273-317). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Cupples, L. (2002). The structural characteristics and on-line comprehension of experiencer-verb sentences. *Language and Cognitive Processes*, 17(2), 125-162.
- Dowty, D. (1989). On the semantic content of the notion of "thematic role". In G. Chierchia, B. H. Partee and R. Turner (Eds.), *Properties, Types and Meaning, II* (pp. 69-129). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3), 547-619.
- Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, 715-736.
- Ferreira, F., Christianson, K., and Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, 30(1), 3-20.
- Ferretti, T., McRae, K., and Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role. *Journal of Memory and Language*, 44(4), 516-547.
- Field, A. (2005). *Discovering statistics using SPSS* London, UK: Sage Publications.
- Ford, M. (1983). A method of obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 203-218.
- Frazier, L., and Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4(2), 93-126.
- Frazier, L., and Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291-325.

- Frazier, L., and Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178-210.
- Gennari, S. P., and MacDonald, M. C. (2003). *Linking production and comprehension of relative clauses*. Paper presented at The 17th Annual CUNY Sentence Processing Conference.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1-76.
- Gordon, P. C., Hendrick, R., and Johnson, M. (2001). Memory Interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27(6), 1411-1423.
- Gordon, P. C., Hendrick, R., and Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, 51(1), 97-114.
- Grimshaw, J. B. (1990). *Argument structure*. Cambridge, Mass MIT Press.
- Grodner, D., and Gibson, E. (2005). Consequences of the Serial Nature of Linguistic Input for Sentential Complexity. *Cognitive Science*, 29(2), 261-290.
- Grodner, D., Gibson, E., and Tunstall, S. (2002). Syntactic complexity in ambiguity resolution. *Journal of Memory and Language*, 46(2), 267-295.
- Haegeman, L. (1991). *Introduction to Government and Binding theory*. Oxford: Basil Blackwell.
- Hale, J. (2001) A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of NAACL*, 2, 159-166.
- Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science* 30, 609-642.
- Jackendoff, R. (1990). *Semantic structures*. Cambridge: MIT Press.

- Jaeger, T. F., and Wasow, T. (2005). Processing as a Source of Accessibility Effects on Variation. In *Proceedings of the 31st Annual Meeting of the Berkeley Linguistic Society*. Berkeley, CA: Berkeley Linguistic Society.
- Jurafsky, D. (2003) Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In Bod, R., Hay, J., and Jannedy, S., (Eds.) *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
- Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122-149.
- Kidd, E., Brandt, S., Lieven, E., and Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Journal Language and Cognitive Processes*, in press.
- King, J., and Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30(5), 580-602.
- Levin, B., and Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry*, 17(4), 623-661.
- Levy, R. (2007) Expectation-based syntactic comprehension. *Cognition*, in press.
- Lewis, R. L., and Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, 29(3), 375-419.
- MacDonald, M., Pearlmutter, N., and Seidenberg, M. (1994). Lexical Nature of Syntactic Ambiguity Resolution. *Psychological Review*, 101(4), 676-703.
- MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32(5).
- MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9(2), 157-201.

- MacDonald, M. C. (1999). Distributional information in language comprehension, production, and acquisition: Three puzzles and a moral. In B. MacWhinney (Ed.), *The Emergence of Language*. Mahwah, NJ: Erlbaum.
- MacWhinney, B. (1977). Starting points. *Language*, 53, 152-187.
- MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Language development: Syntax and semantics* (Vol. 1, pp. 73-136). Hillsdale, NJ: Lawrence Erlbaum.
- MacWhinney, B., and Pleh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, 29(2), 95-141.
- Mak, W. M., Vonk, W., and Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, 47(1), 50-68.
- Mak, W. M., Vonk, W., and Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54(4), 466-490.
- McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283-312.
- Mitchell, D. C., Cuetos, F., Corley, M. M. B., and Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, 24(6), 469-488.
- O'Grady, W., Archibald, J., Aronoff, M., and Rees-Miller, J. (2001). *Contemporary Linguistics: An Introduction* (4th ed. ed.). Bedford: St. Martin's Press.
- Pearlmutter, N. J., and MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, 34(4), 521-542.

- Pickering, M. J., and Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(4), 940-961.
- Race, D. S., and MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In Richard Alterman and D. Kirsh (Eds.), *25th Annual Meeting of the Cognitive Science Society* (pp. 946-951). Mahwah, N.J. : Lawrence Erlbaum.
- Real, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, 57, 1-23.
- Roland, D., Dick, F., and Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, in press.
- Saffran, E. M., and Schwartz, M. F. (1994). Impairments of sentence comprehension. *Philosophical Transactions of the Royal Society: Biological Sciences*, 346, 47-53.
- Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior*, 13(3), 272-281.
- Spivey-Knowlton, M., and Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55(3), 227-267.
- Tabor, W., and Tanenhaus, M. K. (1999). Dynamical models of sentence processing. *Cognitive Science*, 23(4), 491-515.
- Tanenhaus, M., Carlson, G., and Trueswell, J. (1989). The role of thematic structure in interpretation and parsing. *Language and Cognitive Processes*, 4(3-4), SI211-SI234.
- Tanenhaus, M. K., and Trueswell, J. C. (1995). Sentence comprehension. In J. Miller and P. D. Eimas (Eds.), *Speech, language, and communication* (Vol. xviii, pp. 217-262). San Diego, CA, US: Academic Press.

- Traxler, M. J., Morris, R. K., and Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, 47(1), 69-90.
- Traxler, M. J., and Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*. Vol., 35(3), 454-475.
- Traxler, M. J., Williams, R. S., Blozis, S. A., and Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language*, 53(2), 204-224.
- Trueswell, J. C., and Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes*, 6(4), 303-338.
- Trueswell, J. C., Tanenhaus, M. K., and Garnsey, S. M. (1994). Semantic Influences on Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution. *Journal of Memory and Language*, 33, 285-318.
- Trueswell, J. C., Tanenhaus, M. K., and Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden paths. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 528-553.
- Van Dyke, J. A., and Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, 49(3), 285-316.
- Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33(2), 407-330.

- Wanner, E., and Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan and G. A. Miller (Eds.), *Linguistic Theory and psychological reality* (pp. 119-161). Cambridge, MA: MIT Press.
- Warren, T., and Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, 85(1), 79-112.
- Waters, G., Caplan, D., and Hildebrandt, N. (1987). *Working memory and written sentence comprehension*. Coltheart, Max (Ed). (1987). Attention and performance 12: The psychology of reading. (pp. 531-555).
- Waters, G. S., and Caplan, D. (1996a). Processing resource capacity and the comprehension of garden path sentences. *Memory and Cognition* 24(3), 342-355.
- Waters, G. S., and Caplan, D. (1996b). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, 103(4), 761-772.
- Wells, J. B., Christiansen, M. H., and Race, D. S. (2006). Experience and Sentence Comprehension: Statistical Learning, Working Memory, and Individual Differences. Unpublished article. University of Wisconsin-Madison.

Table 1

Example item from which completion prompts were obtained in study 1 and stimulus item in reading comprehension Study 2

<i>Sentence</i>	<i>RC condition</i>
The director that the movie pleased had received a prize.	Active – Animate head
The movie that the director watched had received a prize.	Active – Inanimate head
The director that was pleased by the movie had received a prize.	Passive – Animate head
The movie that was watched by the director had received a prize.	Passive – Inanimate head

Table 2

Relative clause positions at which completions in Study 1 were elicited

<i>RC type</i>	<i>Relative clause prompts</i>			<i>Head-noun</i>
	Position 1: <i>The-N-that</i>	Position 2: <i>The-N-that-the</i>	Position 3: <i>The-N-that-the-N</i>	
<i>Active</i>	The director that ...	The director that the...	The director that the movie...	Animate
	The movie that ...	The movie that the...	The movie that the director...	Inanimate
<i>Passive</i>	Position 1: <i>The-N-that-was-V</i>			
	The director that was pleased			Animate
	The movie that was watched			Inanimate

Table 3
Results of completion task (Study 1) with fragment prompts at various relative clause positions

<i>RC position</i>	<i>Head noun</i>	<i>Roles for NP-that-...NP</i>	<i>RC structure</i>	<i>Prop.</i>	<i>Examples</i>
<i>The-N-that</i>	<i>Animate</i>	Agent-(Theme)	(aux)-V-(NP)	.39	The musician that was playing (the piano)
		Agent-(Goal/Location)	V-(NP/PP)	.16	The child that played at the park/listened to me
		Experiencer-(Theme)	V-(NP)	.25	The contestant that won (the prize)
		Patient-Agent/Theme-Exp.	NP-V-(PP)	.15	The psychologist that I hired/I saw
		Patient-(Agent)	aux-V-(PP)	.05	The psychologist that was fired
	<i>Inanimate</i>	Theme-(Location)	VP	.35	The movie that was playing/was made
		Theme-(Agent/Experiencer)	NP-V-(P)	.65	The article that I read/I saw
<i>The-N-that-the</i>	<i>Animate</i>	Goal-Agent	NP-V-P	.10	The child that the teacher talked to
		Patient-Agent	NP-V-(P)	.50	The student that the teacher suspended
		Theme-Experiencer	NP-V	.31	The teacher that the class liked
	<i>Inanimate</i>	Theme-Agent	NP-V	.63	The grenade that soldier threw
		Theme-Experiencer	NP-V	.18	The incident that the man saw
		Location-Agent	NP-V-P	.07	The school the teacher work/was at

Table 3 (Continued)

<i>The-N-that-the-N</i>	<i>Animate</i>	Agent-Theme	(aux)-V-P	.24	The soldier that the grenade was thrown by
		Experiencer-Theme	V-(P)	.24	The employee that the incident helped/affected
		Goal/Location-theme	(aux)-V-(P)	.22	The soldier that the grenade was given to/went to
		Theme-Theme	V/V-P	.26	The child that the story was about/described
	<i>Inanimate</i>	Theme-Agent	V-(P)	.80	The play the actor wrote/performed in
		Theme-Experiencer	(aux)-V-(P)	.19	The play the actor knew/was involved in

Note: *Exp.* stands for *Experiencer*. The thematic roles are listed in the order of the relative clause nouns: the first role applies to the head noun whereas the second role applies to the noun within the relative clause.

Table 4

Results of completion task (Study 1) in passives at position *The-N-that-was-Ved*

<i>Head noun</i>	<i>Role for by-phrase</i>	<i>Proportion</i>	<i>Examples</i>
<i>Animate</i>	Location	0.03	The plumber that was injured on the job
	Theme	0.28	The musician that was terrified of spiders
	None	0.69	The musician that was terrified
<i>Inanimate</i>	Agent	0.19	The article that was written by the journalist
	Goal/Location	0.09	The incident that was reported to the police/on TV
	None	0.70	The incident that was reported
	Other	0.02	

Table 5

Proportions of correct responses per Relative Clause condition in Comprehension Study 2

<i>RC condition</i>	<i>Mean</i>	<i>SD</i>
Active – Animate head	.69	.25
Active – Inanimate head	.84	.20
Passive – Animate head	.81	.18
Passive – Inanimate head	.84	.16

Table 6

Analysis of variance results at different relative clause word position in Study 2

Region	Effect	By participants			By items			Min F'		
		<i>df</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>minF'</i>	<i>p</i>
Object relative clause noun (e.g., <i>movie</i>)	<i>Animacy</i>	1, 62	15.46	.0002	1, 28	7.80	.009	1, 57	5.18	.03
	<i>Voice</i>	1, 62	7.05	.01	1, 28	3.17	.09	1, 54	2.18	.14
Object relative clause verb (e.g., <i>pleased</i>)	<i>Animacy</i>	1, 62	6.43	.01	1, 27	11.01	.003	1, 87	4.05	.05
	<i>Voice</i>	1, 62	60.74	<.0001	2, 27	33.35	<.0001	1, 57	21.52	.0001
Main verb region (e.g. <i>had received a</i>)	<i>Animacy</i>	1, 62	24.70	<.0001	1, 28	22.21	<.0001	1, 74	11.69	.001
	<i>Voice</i>	1, 62	86.24	<.0001	1, 28	152.18	<.0001	1, 89	55.04	.0001
	<i>Word position</i>	2, 124	56.85	<.0001	2, 56	29.91	<.0001	2, 116	19.59	.0001
	<i>Voice × position</i>	2, 124	15.79	<.0001	2, 56	15.59	<.0001	2, 154	7.84	<.0001
	<i>Voice × animacy</i>	1, 62	6.51	.01	1, 28	9.13	.005	1, 86	3.80	.05

Table 7

Correlation matrix of completion proportions and reaction times at various object relative clause positions from Study 3

Completion predictors		Reading Times at sentence positions				
Count Type	Relative clause positions	Relative clause noun	Relative clause verb	Main auxiliary verb	Main verb	Next word
Structure Proportions	<i>N-that</i>	-.38***		-.31*	-.51***	
	<i>N-that-the-N</i>			-.32*	-.46***	-.32*
Thematic proportions	<i>N-that</i>	-.39***			-.50***	
	<i>N-that-the</i>	-.38***		-.39***	-.34*	
	<i>N-that-the-N</i>			-.40***	-.60***	-.34*
Animacy proportions	<i>N-that-the</i>	-.45***			-.32*	
No. of interpretations	<i>N-that-the</i>	.52***				
	<i>N-that-the-N</i>		.31*	.31*	.31*	
No. of verbs	<i>N-that-the-N</i>		.35**			

Note: "***" = $p < .006$; "**" = $p < .008$; "*" = $p < .02$

Figure Caption

Figure 1: Residual reading times per word position in comprehension study 2.

Figure 2: Examples of scatter plots for different predictors at the different word positions in Study 3.

Appendix A

No.	Sentence stimuli	Thematic roles of relative clause verb
1a	The musician that the accident terrified was in the headlines the next day.	Cause-Experiencer
1b	The accident that the musician caused was in the headlines the next day.	Cause-Theme
1c	The musician that was terrified by the accident was in the headlines the next day.	
1d	The accident that was caused by the musician was in the headlines the next day.	
2a	The contestant that the prize delighted had made a big impression on Mary.	Cause-Experiencer
2b	The prize that the contestant misplaced had made a big impression on Mary.	Agent-Theme
2c	The contestant that was delighted by the prize had made a big impression on Mary.	
2d	The prize that was misplaced by the contestant had made a big impression on Mary.	
3a	The soldier that the grenade injured had come from a town in New Jersey.	Cause-Experiencer
3b	The grenade that the soldier carried had come from a factory in Taiwan.	Agent-Theme
3c	The soldier that was injured by the grenade had come from a town in New Jersey.	
3d	The grenade that was carried by the soldier had come from a factory in Taiwan.	
4a	The scientist that the book annoyed was well-known for his environmentalist writings.	Cause-Experiencer
4b	The book that the scientist studied was well-known for its environmentalist arguments.	Agent-Theme
4c	The scientist that was annoyed by the book was well-known for his environmentalist writings.	
4d	The book that was studied by the scientist was well-known for its environmentalist arguments.	
5a	The director that the movie pleased had received a prize at the film festival.	Cause-Experiencer
5b	The movie that the director watched had received a prize at the film festival.	Agent-Theme
5c	The director that was pleased by the movie had received a prize at the film festival.	
5d	The movie that was watched by the director had received a prize at the film festival.	
6a	The student that the school educated had shown excellent academic performance.	Agent-Patient
6b	The school that the student attended had shown excellent academic performance.	Agent-Theme
6c	The student that was educated by the school had shown excellent academic performance.	
6d	The school that was attended by the student had shown excellent academic performance.	
7a	The teacher that the play angered had taken her students out of the theater.	Cause-Experiencer
7b	The play that the teacher watched had taken several months to rehearse.	Agent-Theme
7c	The teacher that was angered by the play had taken her students out of the theater.	
7d	The play that was watched by the teacher had taken several months to rehearse.	
8a	The employee that the incident bothered had misled the investigation of the facts.	Cause-Experiencer
8b	The incident that the employee reported had misled the investigation of the facts.	Agent-Theme
8c	The employee that was bothered by the incident had misled the investigation of the facts.	
8d	The incident that was reported by the employee had misled the investigation of the facts.	
9a	The plumber that the wrench injured has stayed in the bathroom for a long time.	Cause-Experiencer
9b	The wrench that the plumber dropped has stayed in the bathroom for a long time.	Experiencer-Theme
9c	The plumber that was injured by the wrench has stayed in the bathroom for a long time.	
9d	The wrench that was dropped by the plumber has stayed in the bathroom for a long time.	
10a	The banker that the loan worried had created a problem for the mayor.	Cause-Experiencer
10b	The loan that the banker refused had created a problem for the mayor.	Agent-Theme
10c	The banker that was worried by the loan had created a problem for the mayor.	
10d	The loan that was refused by the banker had created a problem for the mayor.	
11a	The lawyer that the trial confused was covered in the national television.	Cause-Experiencer
11b	The trial that the lawyer reviewed was covered in the national television.	Agent-Theme
11c	The lawyer that was confused by the trial was covered in the national television.	
11d	The trial that was reviewed by the lawyer was covered in the national television.	
12a	The psychologist that the notes annoyed had gotten lost somewhere in the basement.	Cause-Experiencer
12b	The notes that the psychologist printed had gotten lost somewhere in the basement.	Agent-Theme
12c	The psychologist that was annoyed by the notes had gotten lost somewhere in the basement.	
12d	The notes that were printed by the psychologist had gotten lost somewhere in the basement.	
13a	The child that the story scared was from a small fishing village.	Cause-Experiencer
13b	The story that the child told was about a small fishing village.	Agent-Theme
13c	The child that was scared by the story was from a small fishing village.	
13d	The story that was told by the child was about a small fishing village.	
14a	The golfer that the game excited was ignored by most sportswriters.	Cause-Experiencer
14b	The game that the golfer mastered was ignored by most sportswriters.	Agent-Theme
14c	The golfer that was excited by the game was ignored by most sportswriters.	
14d	The game that was mastered by the golfer was ignored by most sportswriters.	

15a	The salesman that the product excited was mentioned in the newsletter.	Cause-Experiencer
15b	The product that the salesman examined was mentioned in the newsletter.	Agent-Theme
15c	The salesman that was excited by the product was mentioned in the newsletter.	
15d	The product that was examined by the salesman was mentioned in the newsletter.	
16a	The fireman that the fire burned had caused trouble at the fire station.	Cause-Experiencer
16b	The fire that the fireman fought had caused serious damage to the house.	Agent-Theme
16c	The fireman that was burned by the fire had caused trouble at the fire station.	
16d	The fire that was fought by the fireman had caused serious damage to the house.	
17a	The fish that the lure attracted had been very close to the boat.	Cause-Experiencer
17b	The lure that the fish attacked had been very close to the boat.	Agent-Theme
17c	The fish that was attracted by the lure had been very close to the boat.	
17d	The lure that was attacked by the fish had been very close to the boat.	
18a	The farmer that the tractor impressed had arrived at the store late last night.	Cause-Experiencer
18b	The tractor that the farmer purchased had arrived at the store late last night.	Agent-Theme
18c	The farmer that was impressed by the tractor had arrived at the store late last night.	
18d	The tractor that was purchased by the farmer had arrived at the store late last night.	
19a	The gardener that the plants pleased has been at the house for several years.	Cause-Experiencer
19b	The plants that the gardener trimmed have been at the house for several years.	Agent-Theme
19c	The gardener that was pleased by the plants has been at the house for several years.	
19d	The plants that were trimmed by the gardener have been at the house for several years.	
20a	The pilot that the plane worried had worked at another company for four years.	Cause-Experiencer
20b	The plane that the pilot crashed had worked without maintenance for four years.	Experiencer-Theme
20c	The pilot that was worried by the plane had worked at another company for four years.	
20d	The plane that was crashed by the pilot had worked without maintenance for four years.	
21a	The executive that the wine relaxed was from a small French village.	Cause-Experiencer
21b	The wine that the executive drank was from a small French village.	Agent-Theme
21c	The executive that was relaxed by the wine was from a small French village.	
21d	The wine that was drunk by the executive was from a small French village.	
22a	The actor that the play delighted had won the first prize at the award dinner.	Cause-Experiencer
22b	The play that the actor rehearsed had won the first prize at the award dinner.	Agent-Theme
22c	The actor that was delighted by the play had won the first prize at the award	
22d	The play that was rehearsed by the actor had won the first prize at the award	
23a	The student that the instrument frustrated had been around for a few months.	Cause-Experiencer
23b	The instrument that the student practiced had been around for a few months.	Agent-Theme
23c	The student that was frustrated by the instrument had been around for a few months.	
23d	The instrument that was practiced by the student had been around for a few months.	
24a	The spy that the message alarmed had come from out of the country.	Cause-Experiencer
24b	The message that the spy encoded had come from out of the country.	Agent-Theme
24c	The spy that was alarmed by the message had come from out of the country.	
24d	The message that was encoded by the spy had come from out of the country.	
25a	The journalist that the article bothered was part of a big scandal.	Cause-Experiencer
25b	The article that the journalist composed was part of a big scandal.	Agent-Theme
25c	The journalist that was bothered by the article was part of a big scandal.	
25d	The article that was composed by the journalist was part of a big scandal.	
26a	The minister that the meal satisfied had been invited to the best restaurant in town.	Cause-Experiencer
26b	The meal that the minister consumed had been made at the best restaurant in town.	Agent-Theme
26c	The minister that was satisfied by the meal had been invited to the best restaurant in town.	
26d	The meal that was consumed by the minister had been made at the best restaurant in town.	
27a	The woman that the jewelry dazzled had been taken to the back room.	Cause-Experiencer
27b	The jewelry that the woman coveted had been taken to the back room.	Experiencer-Theme
27c	The woman that was dazzled by the jewelry had been taken to the back room.	
27d	The jewelry that was coveted by the woman had been taken to the back room.	
28a	The dieter that the dessert tempted was really not very healthy.	Cause-Experiencer
28b	The dessert that the dieter desired was really not very healthy.	Experiencer-Theme
28c	The dieter that was tempted by the dessert was really not very healthy.	
28d	The dessert that was desired by the dieter was really not very healthy.	

Note: the order of the thematic roles reflects the order of arguments in typical subject-verb sentences and not the order of the nouns in the relative clause.

Appendix B

Prompt	Relative clause completions	Thematic Roles coding	Structure coding	Completions proportions					
				Active roles	Active structure	Animacy	Passive roles	Passive structure	Verb count
The soldier that	fought in the war	agent-location	<i>v-pp</i>	0.29	0.14		0.14	0.57	5
	left three years ago	agent-location	<i>v-pp</i>						
	died in combat	experiencer	<i>v-pp</i>						
	died	experiencer	<i>v</i>						
	died	experiencer	<i>v</i>						
	was killed	patient-(agent)	<i>aux-v</i>						
	I met	patient-agent	<i>np-v</i>						
The soldier that the	blast killed	experiencer-cause	<i>np-v</i>	0.67	0.67	0.44			
	president talked to	goal-agent	<i>np-v-p</i>						
	general was talking to	goal-agent	<i>np-aux-v-p</i>						
	woman met	patient-agent	<i>np-v</i>						
	regiment left behind	patient-agent	<i>np-v-adv</i>						
	Vietcong shot	patient-agent	<i>np-v</i>						
	boy saw	theme-experiencer	<i>np-v</i>						
	story told about	theme-theme	<i>np-v-p</i>						
The soldier that the grenade	paper showed	theme-theme	<i>np-v</i>						
	was thrown by	agent-theme	<i>aux-v-p</i>	0.63	0.5				4
	blew up	experiencer-cause	<i>v-p</i>						
	hit	experiencer-theme	<i>v</i>						
	hit	experiencer-theme	<i>v</i>						
	hit	experiencer-theme	<i>v</i>						
	hit	experiencer-theme	<i>v</i>						
	landed near	location-theme	<i>v-p</i>						
The grenade that	was thrown near	location-theme	<i>aux-v-p</i>						
	went off	theme	<i>v</i>	0.29	0.29		0.29	0	4
	exploded	theme	<i>v</i>						
	blew up	theme	<i>v</i>						
	exploded	theme	<i>v</i>						
	exploded	theme	<i>v</i>						
	the soldier threw	theme-agent	<i>np-v</i>						
	the enemy threw at us	theme-agent-goal	<i>np-v-pp</i>						

Appendix B (continued)

The grenade that the	woman threw	theme-agent	<i>np-v</i>	1	0.83	1
	solider pulled	theme-agent	<i>np-v</i>			
	officer threw	theme-agent	<i>np-v</i>			
	soldier threw	theme-agent	<i>np-v</i>			
	soldier threw	theme-agent	<i>np-v</i>			
	soldier threw at us	theme-agent-goal	<i>np-v-pp</i>			
The grenade that the soldier	threw	theme-agent	<i>v</i>	1	1	4
	threw	theme-agent	<i>v</i>			
	threw	theme-agent	<i>v</i>			
	threw	theme-agent	<i>v</i>			
	threw	theme-agent	<i>v</i>			
	had	theme-experiencer	<i>v</i>			
	lost	theme-experiencer	<i>v</i>			
	heard	theme-experiencer	<i>v</i>			

Note: Thematic roles are listed in the order of the relative clause nouns: the first role applies to the head noun whereas the second role applies to the noun within the relative clause (if any).



